DUKE
THE FUQUA
SCHOOL
OF BUSINESS

ECONOMIC EFFICIENCY OF ENFORCEMENT AND ENFORCEMENT-RELATED MONITORING: PROJECT PERIOD ONE REPORT

by
W. Kip Viscusi
Project Director
Wesley A. Magat
Associate Researcher

DUKE

THE FUQUA SCHOOL OF BUSINESS

ECONOMIC EFFICIENCY OF ENFORCEMENT AND ENFORCEMENT-RELATED MONITORING: PROJECT PERIOD ONE REPORT

by

W. Kip Viscusi
Project Director

Wesley A. Magat
Associate Researcher

Center for the Study of Business Regulation

Co-sponsored by the Department of Economics
Institute of Policy Sciences and Public Affairs.
School of Forestry and Environmental Studies, and School of Law

The Fuqua School of Business
Duke University
Durham. North Carolina 27706

ECONOMIC EFFICIENCY OF ENFORCEMENT AND ENFORCEMENT-RELATED MONITORING: PROJECT PERIOD ONE REPORT

by

W. Kip Viscusi, Project Director

Wesley A. Magat, Associate Researcher

CR-811902-01-0

Project Officer

Alan Carlin

U.S. Environmental Protection Agency

Washington, D.C. 20406

March, 1986

Center for the Study of Business Regulation

Fuqua School of Business

Duke University

Durham, North Carolina

DISCLAIMER

Table of Contents

Executive Summary

Chapter 1

An Empirical Analysis of the Effects of EPA Inspections on the Pulp and Paper Industry

1. EPA's regulations are among the most far reaching policies in the risk and environmental field, but there have been few, if any, statistical attempts to analyze the effectiveness of EPA enforcement policies. This chapter reports on a pilot empirical study of the agency's enforcement activities.

2. The primary purpose of this initial phase of the empirical study was to access EPA's PCS data base and to reformat the data in a manner that allows for analysis of enforcement effects upon discharge levels. This task was successfully completed, but it required most of the effort devoted to the empirical portion of Project Period One because the data base was so difficult to use.

3. We also began the task of analyzing the effects of EPA enforcement activities on reductions in pollutant discharges into surrounding waterways by pulp and paper plants. Because the only enforcement actions with relatively complete data in the PCS data base are plant inspections of BOD standards, we limited our attention to the effects of inspections on BOD discharges, leaving the analysis of other

enforcement actions, such as administrative orders and warning letters, to a subsequent study.

4. The second purpose of the pilot empirical study was to analyze the data extensively enough to determine whether EPA inspections, and the associated follow-up activities, did result in subsequent reductions in the discharge of BOD by inspected pulp and paper plants. We successfully demonstrated that EPA inspections do lead to effluent reductions, thus providing a strong justification to study the EPA enforcement process in much more detail in a subsequent study.

5. The initial analysis in this study indicates that pulp and paper plants found to be significantly out of compliance reduce their BOD discharge levels by an average of over 30 percent in the first six months after an inspection. Not surprisingly, sources in compliance and sources not as seriously out of compliance reduce their discharge levels less in response to an inspection. The initial analysis yielded mixed results about whether the discharge reductions are larger after the first six months following an inspection or not, and subsequent work using alternative and more complex models will be needed to resolve this issue.

6. Regression analysis of the data suggests significant differences in the effectiveness of inspections across EPA regions. Because of their policy significance, this initial result justifies further exploration of the magnitudes of these differences and the reasons why they occur. The regression analyses showed no systematic increase or decrease in the effectiveness of inspections over the 1982-1985 study period.

7. The next step in the analysis of the data will be to build an econometric model of the enforcement process that describes both how inspections are scheduled and how those inspections affect the sources' effluent discharge levels. Estimates derived from this model will describe the relationship between inspections scheduled and factors such as previous discharge measurements, the failure of a source to submit Daily Monitoring Reports, the extent of non-compliance, and previous administrative actions. We also plan to examine how the estimates from the model vary across regions, by types of plants, over calendar time, by whether the inspections were state-run or carried out by EPA, and by the type of inspection.

Chapter 2

Review of the Literature on the Enforcement of EPA Standards.

1. The literature on the enforcement of EPA standards is in these areas: (i) effluent changes and standards; (ii) optimal enforcement of regulations; and (iii) permits, licenses, and related policies.

2. Although illustrative EPA data are often provided, there has been no detailed empirical analysis of EPA enforcement policies and their role of EPA penalties. The literature is dominated by what might be termed applied conceptual analyses of economic externality problems.

3. The marginal benefit-marginal cost principles for setting efficient standards can be replicated by setting an appropriate penalty related to the marginal costs generated by pollution.

4. These principles have been refined to take into account uncertainty and the impact of standards on technological innovation.

5. The literature on the conceptual basis for enforcing EPA regulations likewise has a standard orientation, whereas the literature on permits and licenses is concerned not only with the amount of pollution but with the role of pollution pricing in setting optimal levels of pollution. A missing

link in the literature is the role of EPA penalties within the context of a standards enforcement program.

6. A practical barrier to penalty schemes based on the extent of noncompliance as compared with enforcement of standards that must ascertain whether or not a firm is in compliance is that the informational requirements may be greater.

Chapter 3

Procedures for Setting Optimal Penalties for Environmental Regulation

1. EPA enforcement sanctions range in stringency from telephone calls and letters to civil court cases that can lead to civil and criminal sanctions.

2. EPA's measure of the economic gains to the firm from noncompliance (or what EPA calls a "benefits" measure) is the principal penalty approach under the Uniform Civil Penalty Policy.

3. Compliance with EPA air and water pollution regulations appears to be relatively high, but these results are based on self-reported compliance status and neglect the potentially large infractions by those out of compliance.

4. EPA has less control over assessed penalties than do agencies such as OSHA. For example, under the Federal Water Pollution Control Act Amendments EPA can influence penalty levels through its settlement actions and its proposals for court imposed penalties. Under the Clean Air Act EPA has more authority to impose penalties administratively.

5. EPA has the leeway to incorporate both the pollution costs to society and the economic gains to the firm from non-compliance when setting or proposing penalties under the FWPCA. Both factors appear in the published Penalty Policy for water pollution violations, and for air pollution violations the only major constraint is that EPA must impose penalties at least as large as the economic gains to the firm from non-compliance.

6. The setting of EPA penalties is complicated by the inability to monitor with precision the firm's compliance status, the benefits to society of pollution reduction, and the cost savings to the firm from noncompliance.

7. If standards are set at a level where marginal benefits are below marginal costs--an approach that EPA appears to have adopted and which the report concludes may be desirable in situations of regulatory uncertainty--then EPA's benefits-of-noncompliance approach will achieve the desired pollution

control at less cost than any penalty schedule based on a constant price per unit of pollution.

8. By coupling an overly stringent standard with efficient penalties based on the marginal benefits to society of pollution control, then EPA can hedge against the risk that the standard should be tightened and promote efficient levels of pollution control.

9. The relative desirability of using penalties based on the economic gains to the firm from noncompliance as compared with the benefits to society of pollution control hinges on whether EPA's policy objective is full compliance with stated standards or efficient degrees of noncompliance. Since the level of the standard to be set can differ for the two approaches, the issue is not whether EPA wishes to take a strict pollution control policy but how it wishes to approach the pollution control issue.

Chapter 4

The Appropriate Regulation of Involuntary and Hidden Risks

1. Many of the risks regulated by EPA are hidden and/or involuntary risks for which there is no explicit or implicit compensation to the party bearing their risk.

2. Individuals bearing risks involuntarily will value the welfare loss imposed by the risk by more than those who choose to bear the risk voluntarily since (i) the mix of

people who self-select into risky pursuits is different and (ii) the absence of a transaction and compensation for the risk may be consequential.

3. Several concepts for appropriate compensation of victims of involuntary risks can be distinguished: (i) the torts concept of making the victim whole; (ii) providing an efficient level of insurance: and (iii) making the victim whole in terms of lifetime risk.

4. Risks that are hidden tend to be associated with both involuntary imposition and difficult enforcement problems stemming from their clandestine nature.

5. Other things being equal, involuntary risks will not only be associated with a larger-gap between their observed level and their optimal level, but the optimal risk level will be lower as well. EPA should regulate involuntary risks more stringently.

Chapter One

An Empirical Analysis of the Effects of EPA Inspections on the Pulp and Paper Industry

by

Wesley A. Magat

and

W. Kip Viscusi

I. Introduction

Despite the existence of a sizable literature on the enforcement of environmental regulations, there have been few, if any statistical attempts to analyze the effectiveness of EPA enforcement policies. This failure to provide an empirical assessment of EPA's regulatory impacts is particularly striking in view of the burgeoning literature on the effects of auto safety regulation, product safety regulation, and occupational safety and health standards.[1] This comparative inattention to the effects of EPA policies does not stem from any lack of policy importance of EPA efforts. Indeed, by almost any standard EPA regulations are among the most potentially consequential policies in the risk and environmental field.

The potential efficacy of these regulations has not, however, been subjected to systematic scrutiny. There are, of course, occasional studies that point to crude barometers of success, such as changes in air quality levels, but the link of these broad measures of environmental quality to regulatory policy has never been made. In some cases there have even been

---

[1]See, for example, Sam Peltzman, "The Effects of Automobile Safety Regulation, "Journal of Political Economy," Vol. 83(1975), pp. 603-9; W. Kip Viscusi, Regulating Consumer Product Safety (Washington: American Enterprise Institute, 1984); and W. Kip Viscusi, Risk by Choice: Regulating Health and Safety in the Workplace (Cambridge: Harvard University Press, 1983).

major challenges to the meaningfulness of the air quality data itself.

It is this relative inadequacy in environmental data that perhaps most accounts for the lack of research attention devoted to EPA's impact. Data quality problems differ, however, by program type. In the water quality area, which we selected for our study, the underlying pollution measurement data are reasonably accurate, but the structure of the environmental impact data is quite complex. There is, for example, no environmental analog of the data on the rate of lost workday accidents by industry published by the U.S. Bureau of Labor Statistics. Environmental pollutants are multiple and cannot be readily subsumed into a single aggregate measure. Any meaningful empirical assessment consequently must be at a fine level of empirical detail.

To achieve this refinement, we structured the scope of our empirical study quite specifically. In particular, we constructed a set of data on discharge from pulp and paper plants into surrounding water bodies and the EPA inspections for those plants. The focus of our analysis was to assess the relationship between inspections and subsequent reductions in pollution discharge levels. Besides examining the magnitudes of the effects, we are also interested in finding the time pattern of the lagged effects and the variation in effectiveness of inspections across EPA regions so as to provide measures of the relative efficacy of different EPA policies.

After considering several environmental media, we decided to focus the study on water pollution enforcement actions. While EPA maintains records of plant inspections for air pollution regulations, only for water is it possible to find a complete set of discharge measurements that can be directly linked to inspections by plants. Water is an important environmental asset, and much compliance effort has been devoted to water pollution regulations over a long period of time, making it important to resolve the relationship between inspections and pollution reduction for this particular case.

The study considers enforcement actions directed at the pulp and paper industry. This large industry contains several hundred separate sources of water pollution discharges located in seven out of the ten EPA regions. Because the discharges from the pulp and paper industry have long created serious and highly visible pollution problems, this industry has been a major focus of EPA enforcement efforts. While we analyze only one industry in this initial study, there is no reason to suspect that the relationship between EPA water pollution inspections and compliance is fundamentally different in most other industries. The concentration on one industry also avoids the problem of controlling for differences across industries in the stringency of regulations and differences in technology across broad industry groups.

The organic pollutant, Biological Oxygen Demand (BOD), constitutes the most important traditional pollutant in the pulp and paper industry, and has been the prime focus of EPA regulatory efforts.[2] Also, inspections almost always examine BOD levels, as well as test for other pollutants.[3] The abatement technologies used to control BOD discharges tend to reduce the levels of other pollutants as well, suggesting that BOD is the best single indicator pollutant to serve as a proxy for the relationship between inspections and the decrease in all pollutant levels. As a result, the empirical analysis reported below will focus on the effect of EPA inspections on BOD levels.

It should be emphasized, however, the results presented below are only preliminary. The funding for this portion of the project recognized its exploratory nature. Our objective for Project Period One was to construct an enforcement data base and to ascertain whether fruitful analysis with the data was possible. We have done this, but we have by no means completed such analysis.

II. The Data Set

Most of the data for the analysis came from EPA's PCS data set, an extremely large data set collected to assist with monitoring and enforcement efforts for the water pollution rules. All sources are required to report their discharge levels on a monthly basis. Central management at EPA assesses the regional inspection efforts on the data which the regions enter in this

---

[2]BOD is the standard measure of the organic pollutant content of water.

[3]Most of the inspections listed in the PCS data base described below designated BOD as the primary pollutant being examined in the inspection.

data base. Unfortunately, the data on the other enforcement, efforts, such as warning letters and administrative orders, are incomplete because the regions lack the incentives to regularly enter them in the data base. For this reason, in this pilot phase of the study we narrowed our attention to inspections, while leaving for subsequent work the collection of data on the other enforcement actions and the analysis of their effects on discharge levels.

Figure 1.1 lists the nine files contained in the PCS data base. Our analysis makes most use of the Measurements File and the Inspections File. We examine 137 different sources of pollution located in six of the ten EPA regions.[4] The firms are divided into six different four-digit SIC code categories which represent refined components of the pulp and paper industry (SIC code 26). In this study we do not differentiate between the effects of different types of inspections, other than to distinguish between those carried out by EPA and those done by state enforcement agencies. In subsequent work, we plan to analyze the separate effects of the two main types of EPA inspections, compliance sampling inspections and compliance evaluation inspections. Compliance sampling inspections generally require about 30 workdays of time, while in contrast the less extensive compliance evaluation inspections take only about 3 workdays. Of the 276 inspections from 1982-1985 in the

[4]Regions 7,8, and 9 contain no pulp and paper mills. For reasons we do not yet fully understand, the Region 2 inspections data are missing from the file for the years 1982 through 1985.

data base, 42.75 percent are compliance sampling and 57.25 percent are compliance evaluation.

Although some of the data for some of the regions are for years as far back as 1977, we focus on the 1982-1985 period, with the last measurement entered on March 31, 1985. During this time data was being regularly reported by most regions, and the water pollution rules for the pulp and paper industry did not change. However, there was significant change in the vigor of EPA enforcement efforts over this time period, which we did take into account in our subsequent empirical analysis.

While we found the PCS data base to be a rich and useful data source, it presented substantial programming difficulties. The data are stored in a compact and efficient manner, but their format makes retrieval for analysis quite difficult. Despite the fact that this study considers only one industry, it was necessary to extract the relevant information about it from each of the tapes due to the organization of the data base outlined in Figure 1.1. Unfortunately, linking data across files required us to be extremely careful and the task was complex. An added difficulty was that data observations are organized by monthly measurements, rather than by the source of the discharge. For purposes of analyzing the reductions in discharges after inspection <u>by source</u> we needed to average the measurements for each source within a period, as described in the next section. In summary, the use of the PCS data base required extensive data manipulation in order to reformat it into usable files, and this task occupied most of the attention devoted to the empirical portion of this study.

Besides the data in the PCS files, we extracted additional data about pulp and paper firms from Lockwood's Directory. In particular, we used this supplemental data source to find a measure of the capacity of the plants that we will use in subsequent analysis. EPA's GAGE file and the Industrial Facilities Discharge file allowed us to add to our data base measures of the mean and low flow levels of the rivers into which pulp and paper mills discharged their effluents.

III. Research Tasks and Empirical Results

Our primary research task in Project Period One was to access the PCS data base and to reformat the data in a manner that allows for analysis of enforcement effects upon discharges levels. Although the structure of the data base made this task quite difficult, it has been successfully completed. In addition, our other research task was to answer the question of whether EPA inspections, and the associated follow-up activities (e.g., administrative orders and warning letters), resulted in subsequent reductions in the discharge of BOD by the inspected sources. In demonstrating an affirmative answer to this question that reveals a sizable effect of the inspections, we provide strong justification to study the EPA enforcement process in more detail in a subsequent study. We will address this fundamental question through two types of analysis. Section A calculates differences in BOD discharge levels before and after inspections, while Section B takes a regression approach to explaining those differences in discharge levels.

A. Difference Calculations

Table 1.1 presents the mean differences between BOD discharge levels before and after inspections for all pulp and paper source measurements falling into three categories: (i) measurements O-6 months after an inspections; (ii) measurements 7-12 months after their last inspection; and (iii) 13-18 month measurements. In principle, all sources are required to report their discharge levels monthly, but in practice there are often months with missing reports in the data base. Usually, this missing data occurred because the sources failed to report their discharge levels, rather than the failure of the regions to enter the reported measurements into the data base. There is consequently a self-selection bias affecting the distribution of firms across the report categories. Firms that fail to file regular reports will differ from those that do.

For sources with more than one measurement in a six-month period, we averaged the measurements over the period in order that, for each inspection, each source have associated with it only one measurement for purposes of calculating the mean across sources. As well, for any row in the table, such as the first O-6 lag period row, sources with more than one inspection during the 1982-1985 study period contribute one observation for <u>each</u>

inspection that was followed within six months by discharge measurements.[5]

In addition, because the Inspections File often contains missing entries for the measurement variable, we were forced to approximate the measurement at the time of the inspection by the measurement in the Measurements File immediately prior to the inspection date. The mean difference in Table 1.1 then measures the mean (across all sources with at least one measurement in the lagged post-inspection period) of the difference between the measurement just prior to the inspection and the average measurement during the lagged period. Negative values indicate reductions in BOD discharge levels as a consequence of the inspection and all follow-up activity.

This last qualification is important because the firms reacted both to the inspection and to any subsequent actions, such as administrative orders, telephone calls, and warning letters. In future work we plan to seek alternative sources of data on these other enforcement action, besides inspections, and separate their effects on discharge reduction from the effect of the inspection per se. Note that the sample sizes (N) decrease as the lagged period increases in length, primarily due to the

---

[5] In our proposed follow-up study we will estimate the impact on the effectiveness of an inspection of whether or not it closely followed a previous inspection. Also, in the subsequent study we plan to organize the measurements on the basis of average. measurements within three-month calendar quarters, which are common to all sources, rather than taking the measurement averages over six-month periods defined by each source's date of inspection.

fact many sources are reinspected within a year or 18 months of the last inspection.

The mean difference results in column two present an _initial_ picture of the effectiveness of the pulp and paper inspections. Note that the data pertain to different numbers of sources in different lag periods, and they apply to inspections of sources which range from being in compliance to those designated as in a state of significant non-compliance. Subsequent tables will address the influence of these two characteristics on the results reported in Table 1.1. The results in Table 1.1 show a pattern of increasing effectiveness of the inspections through 18 months Averaged across complying and non-compliance sources, there appears to be little effect in the first six months. Although this result could be explained by the need for more than half a year to install the capital equipment required to bring the source into compliance, as well as to implement the operating changes that also contribute to the lower effluent discharge levels, the results in the first row of Table 1.3 will point to another explanation, namely, that the sizable effects of inspections on significantly non-complying firms are masked by averaging them with the effects on sources either in compliance or non-significantly out of compliance. Over the 7 to 12 month period inspections result in an average of a five and one-half percent decline in discharges, and by the end of 18 months the effluent reduction level for all inspected sources averages seven percent.

In order to further clarify the picture of inspection effects drawn by Table 1.1, we recalculated the statistics in the first three rows of the table restricting attention to those 50 source/inspection combinations followed by at least 18 months without a subsequent inspection (and with measurements every half year following the inspection). For those SO source/inspection combinations, again with some inspections finding conditions of significant non-compliance and other finding lesser non-compliance or full compliance, the time pattern of effects, as well as the magnitude of the effects of inspections, differs from that suggested in Table 1.1. After 6 months these sources reduced their discharge levels about eight percent, after 12 months their discharge levels had been reduced by an average of over sixteen percent, and after 18 months the discharge reduction averaged ten and one-half percent.

Given that this sample of sources ranges from those fully in compliance at the time of inspection to those which were far out of compliance with their permitted discharge levels, these results suggest that inspections and subsequent follow-up activities lead to sizable reductions in discharge levels for those sources found to be out of compliance% There appears to be more effect after 12 months than after 6 months, perhaps because sources have had more time to make capital improvements in abatement equipment. Yet even within 6 months the inspections seem to have produced an immediate effect upon compliance. The fact that the average discharge reduction level is less after 18 months than after 12 months may have been caused by the sources' decisions to ease up on control activities because they had not been inspected in over a year and were feeling less pressure to comply, although the difference between the 12-month and 18-month means could easily be explained by random fluctuations in measurements.

While Table 1.2 provides an adjustment for the potential selection bias problem in comparing the means across rows in Table 1.1 (due to different samples of sources being compared), it does not distinguish the effects of inspections on non-complying firms from the effects of inspections on complying firms. At the expense of further reducing the sample size, Table 1.3 directly addresses this latter issue. It provides the analogous statistics to Table 1.1, but restricted to the sources in significant non-compliance at the time of their inspections.

The most important statistic in this table is the 36 percent mean reduction in discharge levels shown to have occurred within 6 months of inspections for significantly non-complying firms. For these sources an inspection leads to a large reduction in discharges very quickly. The mean percentage reductions levels in the second and third rows of Table 1.3 are more difficult to interpret, both because the sample sizes are so low and because the sample composition varies across the three rows of the table. There is some evidence that if sources significantly out of compliance are not reinspected within a year of their initial inspection, they tend to revert to former patterns of non-compliance, although this conclusion is highly speculative and needs to be examined in more detail in our future work with alternative and more complete models of the enforcement process.

B. Regression Equations

We used regression analysis as an alternative approach to measuring the effects of EPA inspections, and their associated enforcement actions, on effluent discharge levels. The dependent variable in the three equations in Table 1.4 again is the difference in BOD discharge levels before and after inspections for all significantly non-complying sources with up to an 18-month lag between inspections and subsequent measurements. For sources with more than one measurement, the measurements differences were averaged to yield one observation per source for each lag period following each inspection.

The explanatory variables, LAG7-12 and LAG12-18, are dummy variables which take the value one when the lagged period for the source was 7 to 12 months or 13 to 18 months, respectively. Their coefficients measure the difference between the reductions in two lagged periods compared with the reductions in the 0 to 6 months period. The TIME variable measures the number of days between January 1, 1982 and the date that the source's inspection occurred. By including this variable in one of the regression equations, we are able to test whether the inspections became systematically more (or less) effective over time during the 1982-1985, independent of the length of the lag between the inspection and the measurements. The REGION3, REGION4, REGION6, and REGION10 dummy variables take the value of one when the source is located in one of those regions, and zero otherwise. Region 1 was chosen as the hold-out region, meaning that the coefficients of the four REGION dummy variables measure the effectiveness of inspections in each of the four regions relative to the effectiveness of an inspection in Region 1.[6]

Equation 1 reinforces the results in Table 1.3, showing a large effect of inspections within the-first six months and lesser effects after longer periods of time have elapsed. Equation 2 tests whether the effectiveness of inspections varied over the study period. The low t-statistic indicates no significant changes in effectiveness of inspections over the 1982 to 1985 period. In another equation (not reported) we added a time-squared variable to allow for non-linear effects of time, but again found no significant effect.

The third equation provides an example of how the data can be used to measure the difference in effectiveness of inspections in different regions. These results should be regarded as very preliminary given the exploratory nature of the analysis. Although the analysis of this issue is still at an exploratory state, the coefficients of the REGION variables in equation 3 indicate that inspections of significantly non-complying firms in Region 6 are most effective of the five regions in the sample, and that Region l's inspections are the least effective ones. The point estimate indicates that Region 6 inspections produce

---

[6]There are no inspections of sources in significant non-compliance in Region 5 during the 1982-1985 period, although this region did carry out many inspections of pulp and paper firms which were either in compliance or not far enough out of compliance to be designated as significantly out of compliance.

discharge reductions which are 90 percent lower than the discharge reductions subsequent to inspections in Region 1. These coefficients, while highly tentative, suggest that there are significant differences in the effectiveness of inspections across regions, and they illustrate the benefits of continued analysis of the data base to determine more precisely how the effectiveness of regions differ and why they differ.[7]

While we have data on the 4-digit SIC code of each source's industry subcategory, whether the inspection was carried out by the state or by the federal EPA, and the capacity level of the plants, there are not enough degrees of freedom in the data base on significantly non-complying firms to allow more refined analysis of the effects on inspection effectiveness of these three variables. In subsequent work we plan to model the enforcement process in a way that increases the sample size of observations enough to allow explorations of the effects of source characteristics, such as product or process differences (SIC code), production capacity, and whether the inspection was carried out by the state or EPA.

IV. Conclusions and Further Research

This pilot empirical project has accomplished its two major tasks. First, the PCS data base has been rearranged into a form

---

[7]For example, it may be that for an identical plant the inspections in Regions 1 and 6 lead to the same reduction in discharges, but plants in Region 1 differ in size or in the type product produced from those in Region 6. Alternatively, one region may conduct fewer inspections than another region, but each of its inspections is more effective.

that is amenable to statistical analysis of the enforcement process. Second, the initial results based on the data base indicate that the EPA inspections process did cause significant reductions in BOD discharge levels during 1982-1985 period in the pulp and paper industry. These results now justify more refined analysis of the enforcement process of several types.

Our next step will be to build an econometric model of the process that describes how both inspections are scheduled and how those inspections affect the sources' effluent discharge levels. The data will be used to reveal the relationship between inspections scheduled and factors such as previous discharge measurements, failure of a source to submit Daily Monitoring Reports, the extent of non-compliance, and previous administrative actions. We also plan to examine how these relationships vary across regions, types of plants, calendar time trends, and enforcement organization (i.e., state or federal EPA). This model will allow us to include sources in the analysis for any given period which both were and were not inspected, in contrast to the approach taken in this study which focuses on only inspected sources.

The model will also allow us to sort out the relative effectiveness of different types of inspections, such as compliance sampling inspections and compliance evaluation inspections, as well as to explain when the agency decides to use one type of inspection rather than the other one.

Because the Inspections File lists only the primary pollutant examined in the inspections for, and this was usually

BOD in the pulp and paper industry, we focused our attention on this pollutant. In subsequent work we plan to link Total Suspended Solids (TSS) measurements to the Inspections File, which will enable us to carry out a similar analysis of this second important pollutant in the industry.

We also plan to examine the effects on compliance of other enforcement variables besides inspections. As explained earlier in the chapter, the only complete enforcement data in the PCS data base is on inspections. By manually searching through Quarterly Non-Compliance Reports we will be able to identify other state actions, and by matching the GREATS data base to the PCS data base we will be able to add administrative orders to our file for analysis.

Finally, after completing the further analysis outlined above, it would be useful in future project periods to replicate the analysis for additional industries besides pulp and paper to test the representativeness of the pulp and paper results. To the extent that regions target their enforcement efforts towards particular industries, we would expect to find inspection effectiveness across regions to be different from one industry to another. Industries may also differ in the mix of enforcement actions taken to force compliance, their financial ability to afford compliance costs, and their technical capabilities of meeting their effluent standards.

Figure 1.1: PCS Data Base Files

File Number	Name	Description of Contents
1	EAPA040-COMPLI.X	Compliance schedules
2	EAPA040.ENFACT.X	Enforcement actions
3	EAPA040.FACLTY.X	Facility identifications and addresses
4	EAPA040.INSPEC.X	Inspections
5	EAPA040.LIMITS.X	Parameter limits
6	EAPA040.MEASUR.X	Measurements
7	EAPA040.0UTFAL.X	Information relating to pipe descriptions
8	EAPA040.PERMIT.X	Permit events
9	EAPA040.VIOLAT.X	Compliance-violation event codes/ data source codes/ comments

Note: X = REGION01
REGION02
REGION03
REGION04
REGION05
REGION06
REGION07
REGION08
REGION09
REGION10

Table 1.1: Mean percentage differences between BOD discharge levels immediately prior to an inspection and the average BOD discharge levels O-6, 7-12, and 13-18 months after the last inspection, for all inspected sources during 1982-1985[a]

Lag Period	Mean	N	Dev.	Min.	Max.
O-6	0.003	211	0.534	-1.991	1.997
7-12	-0.055	137	0.534	-1.542	1.504
13-18	-0.070	64	0.525	-1.997	1.021

[a] The differences are measured in percentages times l/100. Thus, for example, -0.055 indicates a 5.5% reduction in the discharge level after inspection. The differences in discharge levels were transformed into percentages by dividing them by the average of the pre-and post-inspection discharge levels.

Table 1,2: Mean percentage differences between BOD discharge levels immediately prior to an inspection and the average BOD discharge levels 0-6, 7-12, and 13-18 months after the last inspection, for all inspections followed by at least 18 without another inspection, 1982-1985a

Lag Period	Mean	N	Dev.	Min.	Max
0-6	-0.079	50	0.499	-1.992	0.868
7-12	-0.166	50	0.484	-1.260	0.705
13-18	-0.105	50	0.575	-1.997	1.021

[a] The differences are measured in percentages times l/100. Thus, for example, -0.079 indicates a 7.9% reduction in the discharge level after inspection. The differences in dis-charge levels were transformed into percentages by dividing them by the average of the pre-and post-inspection dis-charge levels.

Table 1.3: Mean percentage differences between BOD discharge levels immediately prior to an inspection and the average BOD discharge levels 0-6, 7-12, and 13-18 months after the last inspection, restricted to sources in significant non-compliance at the time of inspection between 1982 and 1985[a]

Lag Period	Mean	N	Dev.	Min	Max
0-6	-0.363	21	0.421	-1.245	0.949
7-12	-0.089	17	0.522	-1.009	0.833
13-18	0.010	9	0.455	-0.556	0.772

[a]The differences are measured in percentages times l/100. Thus, for example, -0.364 indicates a 36.4% reduction in the discharge level after inspection. The differences in discharge levels were transformed into percentages by dividing them by the average of the pre-and post-inspection discharge levels.

Table 1.4: Regressions on the percentage difference between the BOD discharge immediately prior to inspection and the average BOD discharge levels O-6, 7-12, and 12-18 months after the last inspection, restricted to inspections of sources in significant non-compliance at the time of inspection between 1982 and 1985 (t-statistics in parentheses)[a]

Independent Variable	Equation 1	Equation 2	Equation 3
INTERCEPT	-0.299 (-3.053)	-0.414 (-1.859)	0.0541 (0.334)
LAG 7-12	0.220 (0.146)	0.242 (1.548)	0.265 (1.846)
LAG 13-18	0.309 (0.179)	0.191 (1.333)	0.357 (1.987)
TIME		0.000162 (0.655)	
REGION3			-0.331 (-1.483)
REGION4			-0.533 (-2.368)
REGION6			-0.902 (-3.155)
REGION10			-0.393 (-2.058)
N	46	42	46
R-SQUARE	0.081	0.086	0.287
ADJ R-SQ	0.039	0.015	0.180

[a]The differences are measured in percentages times l/100. The differences in discharge levels were transformed into percentages by dividing them by the average of the pre-and post-inspection discharge levels.

Chapter Two

Review of the Literature
on the Enforcement of EPA Standards

by

W. Kip Viscusi

I. Overview

This literature review consists primarily of an annotated bibliography and summary of the principal articles that have appeared relating to the enforcement of EPA regulations. The topics have been divided into three general areas. The first area considers the basic issue of either setting the penalty or standard-level for violations of EPA policies. The second area considers the optimal enforcement of these regulations, particularly taking into account the fact that not all firms will comply with these regulations voluntarily. The third portion of the review consists of an analysis of more innovative policy alternatives along the lines of permits, licenses, and related policies. These efforts and the literature on them shed some light on the more general issue of setting appropriate penalties for violations of EPA standards.

The first series of articles considers the appropriate setting of standards and penalties. The review begins with a summary of some chapters from the environmental economics book by Baumol and Oates. This chapter addresses some of the traditional concerns in this literature. In particular, standards should be set at a level to equate marginal benefits and marginal costs. Alternatively, one can obtain the same result by setting the penalty level equal to the marginal cost imposed by the externality that is generated. This marginal cost link for setting penalty levels has been emphasized by economists since the time of Pigou.

The next series of articles addresses the role of uncertainty in standard setting. In particular, is it preferable to set standards or penalties in a situation in which the marginal benefits and marginal costs may be uncertain? The answer depends on whether the greater uncertainty is with respect to marginal benefits or marginal costs. If there is relatively precise knowledge of the marginal benefits of environmental control and substantial uncertainty in the marginal cost of compliance, or alternatively substantial heterogeneity in the cost of compliance, then it is optimal to impose some kind of fine or tax rather than to rely on direct standards. The final article in this section addresses the differing impact of standards and effluent charges on technological innovation.

The next section of articles addresses the role of enforcement of EPA regulations and, in particular, the consequences of incomplete compliance. Portions of the book by Russell, Harrington, and Vaughan provide a general review of the literature in the area. One of the principal articles reviewed in this survey is that by Viscusi and Zeckhauser, who consider the consequences of incomplete compliance. They show, for example, that tightening the standard may have undesirable effects since it may increase the noncompliance with the regulation by more than it boosts the level of environmental quality for firms that do choose to comply. A series of other articles regarding incomplete compliance is also included, with a common result in many of these being that when analyzing the marginal costs associated with pollution control one should also take into account the marginal costs of enforcement. Finally, this section of articles concludes with a summary of a series of articles addressing both broad optimal legal enforcement issues as well as very narrow issues pertaining to the enforcement of EPA regulations.

The final section of the literature review discusses permits, licenses, and other innovative policies of this kind. Although the task of setting EPA penalties is not identical to designing a license scheme, many of the same kinds of concerns that have been raised in this extensive literature also pertain to the reliance on penalties. For example, there may be important monitoring costs associated with setting a penalty based on the degree of noncompliance as opposed to simply whether or not a firm is in compliance.

II. Effluent Charges and Standards

1. William Baumol and Wallace Oates, Economics, Environmental Policy, and the Quality of Life (Englewood Cliffs: Prentice-Hall, 1979).

2. Robert Crandall, Controlling Industrial Pollution: The Economics and Politics of Clean Air (Washington: Brookings Institution, 1983), esp. Chapter IV.

3. Michael Spence and Martin Weitzman, "Regulatory Strategies for Pollution Control," in Ann Friedlaender, ed., Approaches to Controlling Air Pollution (Cambridge: MIT Press, 1978).

4. Marc Roberts and Michael Spence, "Effluent Charges and Licenses under Uncertainty," Journal of Public Economics, Vol. 5 (1976).

5. Martin L. Weitzman, "Prices vs. Quantities," Review of Economic Studies (1974).

6. Wesley Magat, "Pollution Control and Technological Advance: A Dynmaic Model of the Firm," Journal of Environmental Economics and Management, Vol. 5 (1978).

William Baumol and Wallace Oates, Economics, Environmental Policy, and the Quality of Life (Englewood Cliffs: Prentice-Hall, 1979).

This book provides a comprehensive introduction to environmental issues including those relating to the setting of penalties and standards. In chapter 15 they explore potential market processes that could be used to promote environmental quality. One approach is to tax environmental damages. This could be done in one of two ways. First, the government could set an appropriate tax to reflect the social damage incurred. A second approach is that the tax could be set so as to lead firms to comply with a particular standard. An alternative approach is to adopt a subsidy for pollution control rather than a tax on pollution, but the subsidy creates somewhat perverse incentives since it encourages people to enter into polluting industries so that they can be subsidized. A final approach is instead of setting a tax to let an effective tax be set through the trading of marketable licenses. There are two ways in which the rights to the licenses could be allocated initially. They could be allocated according to bids on the part of firms, with licenses going to the highest bidder, or else they could be distributed equally among firms and then resold.

In Chapter 16 the authors explore some of the problems associated with the pollution charge approach. They distinguish three principal difficulties, some of which also may pertain to standards as well. First, the magnitude of the effect of the pollution charge on pollution may be uncertain. Second, charges may tend to be an inflexible policy over time. It may be difficult to change the tax. (The authors do not raise the issue of whether it is even more difficult to change standards.) And third, the authors believe that it may be politically difficult to vary the tax level by location, which it would be optimal to do.

In Chapter 17 they explore in detail the merits of effluent charges as compared with subsidies. The authors prefer the effluent tax to avoid encouraging non-polluting industries.

In Chapter 12 they explore the potential advantages of employing direct controls of pollution. They distinguish three circumstances under which they would advocate direct controls. First, direct controls would be desirable if metering emissions is impossible. Second, direct controls are desirable if rapidly changing environmental conditions may arise with the risk of catastrophe and finally, direct controls may be desirable in the case of extremely hazardous pollutants.

In Chapter 20 the authors note that direct controls of pollution also have their difficulties. For example, they are often hard to tailor to particular circumstances at the firm. The permits in particular often may not be set at optimal levels.

In Chapter 20 the authors advocate varying the fee charged for pollution according to the damage that the pollutant does. Overall, they favor a mixed approach which would combine standards and effluent charges.

Robert Crandall, Controlling Industrial Pollution: The Economics and Politics of Clean Air (Washington: Brookings Institution, 1983), esp. Chapter IV.

In this book, Crandall reviews the principles for setting standards and penalties and also provides his own proposal in this area. The basic principles are widely known. Consider a situation where there is a marginal benefit and marginal cost curve for pollution removal. The optimum will be at the intersection of these curves. This optimum can be attained either by setting the standard at this optimal level or by setting a penalty equal to the value of the marginal benefits (and marginal costs) at the intersection. Such a penalty will lead firms to reduce pollution until marginal costs just equal marginal benefits at that point.

One problem that has been explored in the literature is that there is not often certainty with respect to the shape of the marginal benefit and marginal cost curves. This problem, in particular, has been the focus of the work by Spence and Weitzman, which is discussed elsewhere in this literature review. If the government sets standards it will fix the level of pollution that can possibly arise, but it will leave the cost levels uncertain. Alternatively, fees or emissions charges of various kinds fix the cleanup costs, but leave the level of pollution uncertain. Which of these policy approaches is more desirable depends on the nature of the uncertainty as well as on the nature of the benefit and cost functions. In the extreme case where the benefit function is flat and costs are uncertain, it is preferable to use fees rather than standards. The opposite result occurs if the greater uncertainty is with respect to the marginal benefit level.

Crandall summarizes three arguments in favor of the superiority of fees or pollution taxes. First, these policy approaches are desirable in situations where there is uncertainty with respect to the pollution control costs. Second, heterogeneity in the cost of compliance will lead to the desirability of using fees as a more flexible policy approach. Third, fees will eliminate the possibility of firms' reducing pollution when it is inefficient to do so.

The principal opposition to pollution taxes is the oft-stated argument that it is tougher to administer a fee system. Crandall notes that this is often alleged to be a difficulty, but there has been little or no documentation of this point. From a practical perspective, the administrative difficulty of fees is that the government is required to measure a continuum of outcomes rather than simply a cutoff point. More generally, the government requires more information to administer such a system.

The principal novelty of Crandall's analysis is that he points out a major practical disadvantage of fees from the standpoint of their attractiveness to firms. Suppose, for example, that we are in a situation where either with a pollution fee system or a standards system the firm would provide the same level of pollution control. The firm in this situation will prefer standards to fees because the only cost to the firm of complying with the standard will be the cost of controlling the pollution. In contrast, under a fee system, the firm will not only have to pay for the pollution control expenditure but also for the cost of all the pollutants that are not removed through this pollution control effort. In effect, a standard gives the firm some free pollution rights so long as it meets the standard, whereas a fee system does not. These penalties could be very expensive depending on the particular situation.

Crandall's policy proposal is a two-part pollution tax. In effect, each firm would be given free rights to pollute up to a particular level. Above that level the firm would have to pay a pollution fee. This two-part pollution tax consequently would yield the same efficiency properties as would a standard effluent charge system, but it would avoid the unattractiveness to firms of a policy under which it would have to pay for all of the pollution that it causes.

Michael Spence and Martin Weitzman, "Regulatory Strategies for Pollution Control," in Ann Friedlaender, ed., Approaches to Controlling Air Pollution (Cambridge: MIT Press, 1978).

This paper is a policy-oriented summary of the more theoretical papers by Roberts and Spence and by Weitzman. The authors conclude that to run a fully optimal effluent charge policy, policymakers generally would require too much information compared with what is readily available. As a result, in terms of an overall policy they recommend a combined approach utilizing both standards and effluent charges.

The problem of pollution as they view it is one of a missing market, which is also an approach that has been noted earlier by people such as Arrow. By setting effluent charges efficiently, the government agency can produce incentives for efficient clean up of pollution.

The particular focus of their paper is on the role of uncertainty. In particular, they consider uncertainty with respect to the damage of pollution clean-up as well as uncertainty with respect to the costs of pollution clean-up. In terms of its impact, setting pollution standards will fix the level of pollution, but will leave the clean-up costs uncertain. In contrast, setting fees will fix the clean-up costs but will leave the levels of pollution uncertain. Whether standards or fees are preferable depends in large part on whether the uncertainty with respect to the costs or benefits is greater. A much more lucid exposition of these particular points is provided in the book by Robert Crandall, which also provides a graphical exposition of these issues.

Marc Roberts and Michael Spence, "Effluent Charges and Licenses under Uncertainty," Journalof Public Economics, Vol. 5 (1976).

This is a technical article that focuses on the issue of uncertainty with respect to the clean-up costs at the firm. In that situation, should the government employ standards, taxes, or some other policy mix? The authors conclude that in this situation it is preferable to use pollution licenses supplemented by an effluent subsidy and a finite penalty.

The model that the authors consider involves an irreversible investment on the part of firms so that the government can't alter the policy over time as it learns about the particular situation. This certainly may be true of a particular firm, but it does not seem to be representative of how the EPA could respond to new entrants and new facilities. They conclude within the context of their model that the optimal policy involves a policy mix. They would impose effluent charges and restrict emissions through licenses. The combination of these policies is what they favor rather than either policy alone. They note that in the case of a linear damage function that effluent charges are optimal. If, however, marginal damages increase with the level of effluents, licenses will be a superior policy, They also note that the uncertainty with respect to the cost of compliance also will lead to uncertainty with respect to the pollution control benefits.

only have to pay for the pollution control expenditure but also for the cost of all the pollutants that are not removed through this pollution control effort. In effect, a standard gives the firm some free pollution rights so long as it meets the standard, whereas a fee system does not. These penalties could be very expensive depending on the particular situation.

Crandall's policy proposal is a two-part pollution tax. In effect, each firm would be given free rights to pollute up to a particular level. Above that level the firm would have to pay a pollution fee. This two-part pollution tax consequently would yield the same efficiency properties as would a standard effluent charge system, but it would avoid the unattractiveness to firms of a policy under which it would have to pay for all of the pollution that it causes.

Martin L. Weitzman, "Prices vs. Quantities," Review of Economic Studies (1974).

This paper provides a theoretical analysis of the use of price or quantity information in a decentralized economy as a mechanism for appropriate planning. One can view setting standards as equivalent to setting the quantity of pollution. Similarly, one can view setting effluent charges as equivalent to setting the price for pollution. Weitzman analyzes the implications of prices vs. quantities for the situation where both benefits and costs may be uncertain. Depending on the relative uncertainty in the shape of the benefit cost curves, one strategy may dominate the other.

Wesley Magat, "Pollution Control and Technological Advance: A Dynamic Model of the Firm," Journal of Environmental Economics and Management, Vol. 5 (1978).

This paper is concerned with the different impact of taxes and standards on technological innovation. The paper addresses in particular whether a constant standard or a constant tax rate for pollution will be preferable in terms of the incentives that it provides for technological advance. One of the conclusions of the paper is that a rising tax is often needed to promote the same outcome in terms of an effluent level that would occur with a uniform standard. In addition, the usual equivalence between a standard and a penalty in a single period model becomes much more complicated within the context of a multi-period model with technological change. Depending on the time pattern of penalties and standards, the policies may have identical effects or effects on innovation that are difficult to generalize in terms of a simple result, such as that effluent charges dominate standards in terms of the impact on technological innovation.

III. Optimal Enforcement of Regulations

7. Clifford Russell, Winston Harrington, and William Vaughan, Monitoring and Enforcement in Pollution Control, draft book manuscript, Resources for the Future, 1985.

8. W. Kip Viscusi and Richard J. Zeckhauser, "Optimal Standards with Incomplete Enforcement," Public Policy, Vol. 27 (1979).

9. Jon Harford, "Firm Behavior under Imperfectly Enforceable Pollution Standards and Taxes," Journal of Environmental Economics and Management, Vol. 5 (1978).

10. William Drayton, "Economic Law Enforcement," Harvard Environmental Law Review, Vol. 4 (1980).

11. A. Mitchell Polinsky and Steven Shavell, "The Optimal Tradeoff between the Probability and Magnitude of Fines," American Economic Review, Vol. 69 (1979).

12. George Stigler, "The Optimum Enforcement of Laws," Journal of Political Economy, Vol. 78 (1970);

13. Dwight R. Leigh, "The Economics of Enforcing Pollution Taxation," Journal of Environmental Economics, Vol. 11 (1984).

14. Paul Downing and William Watson, "The Economics of Enforcing Air Pollution Controls," Journal of Environmental Economics and Management, Vol. 1 (1974).

15. Steven H. Linder and Mark McBride, "Enforcement Costs and Regulatory Reform: The Agency and Firm Response," Journal of Environmental Economics and Management, Vol. 11 (1974).

16. William Watson and Paul Downing, "Enforcement of Environmental Standards and the Central Limit Theorem," Journal of the American Statistical Association, Vol. 71(1976).

Clifford Russell, Winston Harrington, and William Vaughan, Monitoring and Enforcement in Pollution Control, draft book manuscript, Resources for the Future, 1985.

This book on monitoring begins by questioning the often made assumption that there will be perfect compliance on the part of firms and perfect monitoring on the part of EPA. They note that self-monitoring has been written into the permits of the National Pollution Discharge Elimination System (NPDES) for water pollution and into the New Source Performance Standards for air pollution. They conclude that in many situations, particularly in the major industries, firms have initially complied. In particular, they have often installed equipment to provide for potential compliance with EPA standards. They claim, however, that there is a problem of continuing compliance whereby firms have an ongoing difficulty of meeting discharge limits.

They distinguish different ways in which standards could be set. For example, they could be set based on the concentration of pollutants, the weight of pollutants emitted per unit time, the weight per unit production, the fuel or other input quantity, the percent of pollution removal, or the ambient standards around a particular source.

The authors suggest that legal penalties are often cumbersome and as a result states have set up administrative penalties to provide incentives for compliance. Promoting compliance is, however, a costly undertaking. There are significant monitoring and surveillance and test costs associated with enforcement. In Connecticut penalties were set on the basis of the economic benefit of noncompliance. The overall conclusion of the authors is that there is incomplete compliance and very few sanctions for noncompliance.

The authors provide a review of some of the more significant studies of enforcement, which are provided in the attached table. The most basic model in which compliance is assumed is treated in various literature reviews by, for example, Teitenberg. Two models dealing with the potential noncompliance have been developed by Viscusi and Zeckhauser. The fourth class of models in which compliance is assumed but discharges are stochastic is that developed, for example, by Downing and Watson. Finally, the potential noncompliance model associated with a stochastic noncompliance situation is that developed by Linder and McBride.

The authors suggest that in many of these models it may be appropriate to have the probability of detection of a violation linked to the size of the violation. The authors developed a Markov model of transitions among four states: whether or not the firm is in compliance and whether or not the firm is under surveillance by EPA. In effect, these binary outcomes are crossed leading to a 2x2 format.

Finally, the authors developed a quality control model, The sequence of decisions in their pollution game involve compliance vs. violation, detection vs. no detection, and acceptance of the monitoring report vs. rejection of the reading.

W. Kip Viscusi and Richard J. Zeckhauser, "Optimal Standards with Incomplete Enforcement," Public Policy, Vol. 27 (1979).

This paper analyzes the setting of optimal standards in situations in which compliance is not complete. More specifically, firms may choose not to comply with a particular environmental standard if the costs of compliance exceed the expected cost of noncompliance that will result from the EPA enforcement effort. Rather than analyzing a particular firm, this paper explicitly takes into account the heterogeneity of compliance cost in the industry. The tightness of the environmental standards set by EPA consequently will affect the cutoff point on the distribution of firms that will be in compliance and those that will- not. Even if the government's objective is solely to promote environmental quality, it is shown that tightening the standard may not be advantagous. The difficulty is that a tighter standard will increase the degree to which firms find it desirable not to comply with the standard. This paper considers both the decision to comply before the inspection as well as the decision to comply after an EPA inspection has been taken.

To reduce the problems of noncompliance, the government might wish to explore heterogeneous standards. Although variation of standards on an individual firm basis is infeasible, many profitable distinctions might be drawn. For example, the government could impose different standards on new facilities than on existing facilities. Tighter standards for new facilities are optimal in this situation. As a result, the "new source bias" that EPA policies often display is quite consistent with what is economically efficient. The degree of the new source bias may not, however, necessarily be efficient.

Jon Harford, "Firm Behavior under Imperfectly Enforceable Pollution Standards and Taxes," Journal of Environmental Economics and Management, Vol. 5 (1978).

This paper analyzes the behavior of firms in a situation where the pollution standards or taxes are not completely enforced. As a result, noncompliance is permitted. As one might expect, policies that increase the expected penalty to the firm associated with noncompliance will increase the pollution control investment by the firm. Making the standard more stringent will not necessarily increase the degree of pollution control. Whether or not this will in fact occur depends on how the tightening of the standard interacts with the schedule of penalties faced by the firm for noncompliance. To take an extreme case, for example, if a firm is going to be fined by the same amount for small and large violations, a tight standard may lead to much more noncompliance because the cost of compliance has risen and the penalties associated with noncompliance have not.

The author shows that the marginal cost of pollution reduction by the firm will be equated to a constant pollution tax by the firm. As a result, the degree of noncompliance will be independent of the actual amount of wastes. The author modifies the usual maxim that marginal benefits should be equated to marginal costs by noting that the marginal costs should also include the costs associated with enforcement.

William Drayton, "Economic Law Enforcement," Harvard Environmental Law Review, Vol. 4 (1980).

This article provides a review of the Connecticut enforcement project on which Drayton participated. In particular, in Connecticut a delayed compliance penalty was imposed on noncomplying firms to provide an incentive for compliance. Drayton concludes that increasing the enforcement sanction had beneficial effects in Connecticut.

The Connecticut approach was based on two principles. First, the penalties were based on an economic recapture principle. The object was to eliminate the potential economic gains the firm might have from noncompliance. The second principle is that there would be escalating penalties for noncompliance to increase the incentives to comply once a violation had been discovered. Drayton notes that the approach basing the penalty on the gain to firms rather than on the environmental damage is not likely to be accepted by economists, but that it would be acceptable if there is a strong correlation between the damages inflicted on the environment and the clean-up costs firms must incur.

A. Mitchell Polinsky and Steven Shavell, "The Optimal Tradeoff between the Probability and Magnitude of Fines," American Economic Review, Vol. 69 (1979).

This paper does not focus on environmental issues per se, but rather addresses the more general issue of altering the intensity of the enforcement effort (i.e., the probability of fines) and the level of the penalty. The authors first show the familiar result that if the violators are risk neutral, then the optimal fines are infinite. These will insure that there is compliance with whatever standard must be met. In the case in which there is risk aversion, if the cost of catching violators is sufficiently small it is optimal to catch all violators, as one would expect. Penalties exceeding the external cost that the violators generate may be optimal, but there is a need to mute the extent of the punitive penalty because of the presence of risk aversion. As a result, they conclude that it is often desirable not to set the penalty not too much above what would be dictated on the basis of the external costs generated.

George Stigler, "The Optimum Enforcement of Laws," Journal of Political Economy, Vol. 78 (1970).

Although this paper does not focus on environmental problems, it nevertheless provides at least one practical insight for environmental policy. Stigler emphasizes that policies should be designed to provide for an affordable degree of compliance. In particular, they should attempt to provide marginal deterrence for violations of differing magnitude. For example, if an individual is going to be sent to life inprisonment for either murder or for armed robbery, there is no marginal incentive for the criminal to avoid shooting two potential witnesses. Similarly, in the environmental context if there are no marginal incentives for pollution control, a firm will have no incentive to reduce its pollution at all once it is out of compliance.

Dwight R. Leigh, "The Economics of Enforcing Pollution Taxation," Journal of Environmental Economics, Vol. ll(1984).

This paper analyzes the appropriate pollution tax in a situation in which there are not only costs to the firms of complying with the EPA standard, but there are also costs to EPA of enforcing its regulations. It is this additional enforcement cost component that distinguishes the author's model from those in the literature. The polluters in this model have an incentive to avoid the tax on pollution, and the government must incur costs to increase the likelihood of detection. The firm's choice variables are the pollution rate and the amount of avoidance behavior, and its profits are given by the private benefits of pollution minus the total pollution tax and the costs associated with avoiding the tax. The result that the authors derive is that it will generally be optimal to have a lower tax rate, but more stringent enforcement when costs of avoidance are considered than when it is not considered.

Paul Downing and William Watson, "The Economics of Enforcing Air Pollution Controls," _Journal of Environmental Economics and Management_, Vol. 1 (1974).

This paper analyzes first the equalization of marginal social costs and social benefits of pollution control. It includes in the social cost both the cost to the firm as well as the costs of legal enforcement. The neglect of the enforcement cost would lead to standards that are inefficiently stringent. The authors then develop a model of optimal enforcement policies, which is an extension of Becker's crime model. Firms are assumed to minimize expected costs, where these include the expected cost of noncompliance actions taken against the firm.

The authors then develop a simulation model for six policy scenarios involving various kinds of penalty systems and different assumptions about the flexibility of the technology. These were done for the case of air pollution and, in particular, for opacity standards and fly ash standards. For the scenarios they examine, an effluent fee enforcement scheme would be optimal in controlling fly ash emissions from coal-fired power plants.

Steven H. Linder and Mark McBride, "Enforcement Costs and Regulatory Reform: The Agency and Firm Response," _Journal of Environmental Economics and Management_, Vol. 11 (1974).

This paper considers the optimal enforcement of EPA regulations and, in particular, the response by firms to this enforcement effort. As in other papers of this type, one does not simply equate marginal benefits and marginal costs of pollution control but rather the marginal cost term is comprised of both the marginal cost of pollution reduction as well as the marginal cost of enforcement. The emphasis of this paper is on the optimal testing properties of the EPA enforcement effort. In particular, during any enforcement inspection, a violation might be discovered, a violation that is present may be missed, a violation may be alleged but not be present, and a violation may be found. This formulation is drawn from the standard literature on Type I and Type II errors. As in that literature, there is a trade-off between Type I and Type II errors. The authors explore different ways of reducing these errors by, for example, installing better monitoring equipment.

William Watson and Paul Downing, "Enforcement of Environmental Standards and the Central Limit Theorem," Journal of the American Statistical Association, Vol. 71 (1976).

This paper is an exploration in a particular case study of the stochastic properties of pollution control devices. In particular the authors focus on the fly ash violations that may arise. The authors conclude that even in a situation where a technology is potentially in compliance with the standard, due to random variations in the performance of the pollution control device it may be out of compliance on a stochastic basis.

IV. Permits, Licenses, and Related Policies

17. W. David Montgomery, "Markets in Licenses and Efficient Pollution Control Programs," Journal of Economic Theory, Vol. 5 (1972).

18. Robert Hahn and Roger Noll, "Barriers to Implementing Tradable Air Pollution Permits: Problems of Regulatory Interactions," Yale Journal on Regulation, Vol. 1 (1983).

19. Robert Hahn and Roger Noll, "Designing a Market for Tradable Emissions Permits," in Wesley Magat, ed., Reform of Environmental Regulation (Cambridge: Ballinger Publishing, 1982).

20. Susan Rose-Ackerman, "Effluent Charges: A Critique," Canadian Journal of Economics, Vol. 6 (1973).

21. Thomas Teitenberg, "Transferable Discharge Permits and the Control of Stationary Source Air Pollution: A Survey and Synthesis," Land Economics, Vol. 6, No. 4 1980.

22. Jorge A. de1 Calvo y Gonzalez, "Markets in Air: Problems and Prospects of Controlled Trading," Harvard Environmental Law Review, Vol. 5 (1981).

W. David Montgomery, "Markets in Licenses and Efficient Pollution Control Programs," <u>Journal of Economic Theory</u>, Vol. 5 (1972).

This article is a conceptual analysis of the implications of a pollution pricing scheme as an alternative to standards. Montgomery distinguishes two different approachs. First, the government could adopt a pollution license approach whereby we would impose a fee based on the total amount of pollution generated. This is the approach that Montgomery prefers. The second approach is an emissions license approach where the fee would be based on the rate of pollution. Montgomery believes that the latter policy involves more difficulties because of the differing impact of differing rates of pollution across firms.

In a model in which he explores the implications of the creation of this new market, Montgomery concludes that both licenses approaches can lead to efficient outcomes. In effect, he is fulfilling the missing markets notion developed in an earlier paper by Kenneth Arrow. The principal benefit of this approach is that it establishes on a sound theoretical basis the efficiency properties of the licenses-fee approach to pollution control. In effect, the government can produce an efficient market in this situation by setting up a market where one otherwise did not exist.

Robert Hahn and Roger Noll, "Barriers to Implementing Tradable Air Pollution Permits: Problems of Regulatory Interactions," Yale Journal on Regulation, Vol. 1 (1983).

This paper is a continuation of the exploration of the properties of a market for air pollution permits for sulfer oxide pollutants in the Los Angeles area. The authors continue to be in favor of the permits approach and they note that one potential advantage that arises from it is that tradable permits ease the entry and exit of polluters into an industry as compared with source-specific regulations. In particular, in terms of the dynamics of the way in which the economy develops, the tradable permits approach has substantial efficiency advantages. It creates incentives for innovative pollution-controlling technologies that would not be present under a simple technological standards policy.

Robert Hahn and Roger Noll, "Designing a Market for Tradable Emissions Permits," in Wesley Magat, ed., Reform of Environmental Regulation (Cambridge: Ballinger Publishing, 1982).

This paper focuses on the design of a tradable emissions permit market for sulfer oxides in the Los Angeles area. This is a more radical approach than the controlled trading option that had been explored earlier by EPA. The authors conclude that the emissions permit is a promising and practical policy approach. In particular, they stress that it provides beneficial incentives for promoting environmental quality.

Susan Rose-Ackerman, "Effluent Charges: A Critique," _Canadian Journal of Economics_, Vol. 6 (1973).

This paper provides a review and critique of the effluent charge approach. The author makes the following points with respect to the potential limitations of the effluent charge policy approach: 1) there may be significant costs in terms of the information that the environmental agency must collect, 2) there may be a substantial regulatory lag whereby the effluent charges may not respond to environmental needs, 3) this approach may not be sufficiently well-adapted to the capital intensive nature of pollution investments, 4) it neglects the potential difficulties that might arise if there is collusive behavior on the part of pollution dischargers, 5) there may be discontinuities in the marginal benefits and marginal costs that create difficulties for the policy, and 6) it may be hard to administer such a policy optimally if the marginal damage caused by different polluters differs or if a joint treatment option is more efficient.

Although this is an interesting review of the potential limitations of effluent charges, most of these criticisms apply to equal or greater degrees to the standards policy approach. For example, the problem of regulatory lag would seem to affect standards just as much if not more than effluent charges. It is also noteworthy that apart from largely practical difficulties, the author does not dispute the conceptual superiority of the effluent charge approach.

Thomas Teitenberg, "Transferable Discharge Permits and the Control of Stationary Source Air Pollution: A Survey and Synthesis," Land Economics, Vol. 6, No. 4, 1980.

This paper consists largely of a review of the air pollution permits literature. In terms of providing new insights into appropriate setting of EPA penalties or standards, the paper does not break new ground. However, it does provide a useful survey, should one be interested in the discharge permit issue.

Jorge A. de1 Calvo y Gonzalez, "Markets in Air: Problems and Prospects of Controlled Trading," Harvard Environmental Law Review, Vol. 5 (1981).

This paper is largely a survey of the problems of controlled trading in the air pollution area. It includes a survey of EPA policies that provide for controlled trading. The paper also explores the literature on marketable rights schemes. The paper reviews concepts such as offsets, bubbles, and banking, but does not break any new ground from an economic standpoint.

Chapter Three

Procedures for Setting Optimal Penalties for Environmental Regulation

by

W. Kip Viscusi, Richard Zeckhauser, and Peter Schuck

I. The Nature of EPA Enforcement

A. Standards and the Monitoring of Compliance

The U.S. Environmental Protection Agency (EPA) follows a regulatory approach common to other social regulation agencies. In particular, it sets standards or other constraints on environmentally related actions and enforces these standards through inspections. If firms are out of compliance, it attempts to impose penalites on the violators. Although this report focuses on the role of the penalty policy, it is useful to begin with a brief overview of the regulatory approach, since this defines the context in which the penalty policy operates.

The strategy of setting a standard for pollution above which there is a penalty for noncompliance and below which there is no penalty is but one choice that government regulators could have made. The penalty schedule could, for example, have begun with punitive fines at zero level of pollution or with a fine that rose linearly with the level of pollution (a pollution tax). Moreover, the character of the standards is also quite specific as it is often structured in terms of the firm's technology rather than simply the pollution level.

One can envision several possible rationales for the standards-penalty approach. First, the optimal level of pollution may be substantially above zero so that it is not desirable to eliminate all pollution. Setting a standard below which there is no charge for pollution in effect gives the firm some pollution free, increasing its attractiveness over a straight unit penalty system. Second, standards are desirable when it is possible to monitor the regulated activity. If discharges are readily measurable one can set a standard based on pollution levels, and if such monitoring is costly a technology-based standard may be desirable. A third consideration is that the ability to characterize participants into a limited number of classes enables EPA to set standards for an entire group of firms without too much efficiency loss. Requiring, for example, that all pulp and paper mills install primary treatment facilities establishes an appearance of fairness since all firms face similar requirements.

EPA standards take a variety of different forms. In regulating water pollution, for example, EPA may either set a standard for water pollution control or issue a permit approving a certain level of pollution for a particular firm. Each approach in effect is a standard, but permits pertain to standards that have been based on some specific characteristics of a particular facility, such as whether the facility is a refinery or a pulp and paper mill. In the discussion below, we will not usually distinguish between standards and permits.

For most air and water pollution regulation, the EPA standard pertains to the total mass of pollution discharged. Alternatively, standards may limit the concentration of pollutants in the ambient environment, expressed in terms of either the overall concentration of the pollutant (e.g., units of mass per unit volume) or the relative concentration (parts per million). Much less performance-oriented in character are technology standards, which specify the design of particular kinds of control equipment that must be installed. In other cases, EPA imposes practice standards that specify or prohibit particular operating procedures, such as agricultural practices to minimize non-point water runoff.

Much less stringent in their general character are EPA policies that are information oriented. EPA may lay down informational rules, for example, requiring firms to report releases of hazardous wastes. Finally, EPA has a variety of labeling policies, such as its detailed pesticide label program, which provide information to users.

For each of these standards EPA must ascertain whether a firm is in compliance with the regulation. In many cases, this information is supplied to EPA by the regulated firms and facilities. Water pollution regulations, for example, rely heavily on self-monitoring. Firms monitor their discharges and must file regular discharge monitoring reports so that EPA can verify that they are complying with the terms of the permit. Similarly, firms sometimes file stack emissions tests for air pollution standards and must provide a self-analysis of hazardous wastes.

If the regulated firms provided EPA with complete and accurate information there would be little need for an extensive enforcement staff. In practice, however, field visits are necessary to verify the information provided and to gather additional information not supplied to EPA. Field visits vary in intensity, ranging from walk-through inspections to detailed sampling inspections. It is not always necessary to visit the plant site to acquire information. For example, EPA has installed air quality monitors throughout the country.

Which enforcement office undertakes the inspection or enforcement action depends on the particular situation. EPA has separate enforcement offices for each programmatic area of concern, such as air, water, and hazardous wastes. This separation makes it difficult to identify "bad actors" for several types of pollution.

Unlike the criminal class, where one's overall record is of extraordinary importance, little is kept in the way of an environmental record. EPA might find such information could be particularly useful in targeting its efforts, particularly for a large firm with many locations or for a firm emitting pollutants covered by different EPA policies (e.g., air and water). Although most regulations are enforced by EPA itself, the states sometimes have the enforcement responsibility. For example, the states are responsible for enforcing national ambient air pollution standards for existing sources, whereas performance standards for new sources are enforced by EPA. Similarly, 37 states have assumed responsibility for the water pollution permit programs.[1]

B. Enforcement Sanctions

If a violation of regulations is discovered, EPA can take a variety of steps. Perhaps the least severe is a telephone call or a letter alerting the firm to the violation and urging it to comply with the standard. Warning letters often detail the penalties the firm faces if it does not comply with the standard. A representative of the enforcement office also can undertake a visit to the firm to discuss the action that need be taken. Finally, EPA can issue a notice of violation, a formal notification that a violation has occurred. These notices generally specify a deadline for the firm's response.

More formal administrative remedies include requirements that a firm take corrective action. Administrative orders are subject to the appeals process in the Federal and State Administrative Procedures Acts.[2] If firms violate the administrative order, EPA can pursue action through the Justice Department in the U.S. courts to force compliance. These civil court cases may lead to financial penalties on the firm, as discussed below. Finally, EPA can pursue criminal enforcement for willful and serious violations and for fraudulent reports. Civil court remedies are much more frequent than criminal cases. The number of civil cases referred has ranged from 90 to 270 annually in the 1980s, whereas the annual number of criminal cases has been between 20 and 36 over that period.[3]

In some instances the penalties are set by EPA; usually EPA refers the offense to the Justice Dept. for litigation in the courts. Under EPA's recently issued Uniform Civil Penalty Policy, EPA takes into account two factors when setting penalties.[4] First, it takes into account the economic gain to the firm of delayed compliance. The penalty increases as the amount the firm makes from noncompliance rises. Calculating this amount is often complex because of the various tax provisions affecting capital investments. The second component of the penalty is the gravity of the violation. In effect EPA considers both the benefit to society of compliance with the regulation (i.e., the gravity of the violation) and the costs of compliance to the firm (or what EPA terms the economic benefits of noncompliance). This amounts to double-counting, and it will not generally be desirable simply to add up the levels of benefits and costs when setting the penalty level.

The overall intent of these penalties is to produce compliance with the regulation.[5] In practice, however, the penalties have a more subtle role to play. For example, they can be used to signal to firms that the government is serious about its enforcement effort or they can be used to convey the impression that a generally weak enforcement policy with infrequent inspections and a low probability of detecting violations is quite stringent. As with the enforcement of OSHA standards, EPA tries to avoid shutting down firms.[6] The preferential treatment of the steel industry, which has long been a major air pollution violator, is a notable example. More generally, EPA will not seek penalties beyond the means of the violator; instead the agency will spread the penalty out over time or reduce its magnitude.

If penalties and regulations are set properly, there is no need to consider the viability of the firm after an appropriate penalty, It may not be desirable from an economic efficiency standpoint for firms that are significant polluters to continue in business. Social welfare may be better served by designing and enforcing sound regulatory policies and their penalty structure than by permitting selective noncompliance. Some firms or industries may simply not be viable once they are required to take into account the environmental damages caused by their operations, and they should be allowed to fail.

It is particularly desirable for the marginal firms that are serious polluters to go out of business if they are not viable under a judiciously set standards system. Ideally, a firm should pay for all of the pollution costs it imposes on society so as to establish efficient incentives for entry into the industry. With a standards approach, the firm receives the first units of pollution free before it faces any penalties as a result of noncompliance. Firms that are unable to meet the environmental standards consequently may be well below the threshhold of viability in view of the implicit subsidy associated with standards.

There should, nevertheless, be some reluctance to close firms' operations. In their effort to stress the importance of not permitting inefficient polluters from continuing to operate, economists have neglected the real economic costs of plant closings. Workers will encounter substantial adjustment costs that may go beyond their wages alone to the extent that they have been earning economic rents in their positions. Capital equipment also may have much lower scrap value than its value when integrated within the context of the firm's current technology. One can view these losses as being a transactions cost, establishing a gap between the firm's long-run break-even amount and the level of profitability before it is desirable for EPA to shut the firm down. Since the transactions costs are borne largely by individuals other than the owners of the firm, any accommodation of these concerns must reflect a relaxation of the standard or the associated penalties.

A final aspect of the penalty structure other than its level is the intent. Is EPA trying to promote compliance with a standard, or is it using penalties as a pollution fee? In the latter case, the polluter can choose between complying with the standard and paying the penalty while continuing to pollute. In a sense the firm always has that option, but when penalties are intended to force compliance, they are generally set so high as to swamp the firm's costs of compliance.

Although EPA generally views penalties as tools for ensuring compliance,[7] this need not be the case. For example, policy innovations in the area of tradable air pollution rights in effect enable the firm to pollute for a particular price. Penalties could be structured in the same way. The only difference is that for penalties, EPA fixes the price of pollution, whereas in the tradable permits case it fixes the quantity. Either approach can lead to much the same results by setting the penalty or quantity of tradable rights properly. The main difference is that tradable rights may give firms on the threshold of shutting down more leeway to continue operations if its endowment of pollution rights is sufficient.

C. The Magnitude of EPA Penalties

The enabling legislation for EPA programs limits the level of penalties that can be assessed for violations of standards. These provisions establish penalty ceilings for the monetary penalties that can be assessed through civil or criminal action and for the terms of imprisonment that can be imposed in a criminal action. Table 1 summarizes the pertinent limits for six EPA programs.

Quite large fines are permitted. For civil cases, the daily penalty maximum ranges from $5,000 for violations- of the Safe Drinking Water Act and certain provisions of both the Clean Air Act and the Resource Conservation and Recovery Act to $25,000 for violations of the Toxic Substances Control Act and sections of both the Clean Air Act and The Resource Conservation and Recovery Act. The maximum limits for criminal penalties are usually about one year's imprisonment.

Smaller penalties are associated with federal efforts to enforce job safety regulations. Serious violations of an OSHA standard can be penalized as much as $1,000 a day for noncompliance. This sanction, although small, can be assessed administratively by OSHA without filing a court case. Civil penalties, which are seldom imposed, can lead to fines up to $10,000. A possible rationale for smaller OSHA penalties is that the beneficiary of the regulations--the worker-- is present and can help monitor compliance, decreasing the role that agency enforcement efforts must play.

EPA penalties are potentially larger, but they are more difficult to impose. To the extent that the health and safety effects of environmental regulations are greater than those of workplace standards, larger penalties are appropriate. It is entirely possible, however, that neither agency assesses the ideal absolute level of penalties, in which case a comparison of penalty levels across agencies would not be a good index of the appropriateness of EPA penalty amounts.

The lower penalty limits specified in statutes are seldom binding. EPA often deals with a noncomplying polluter through enforcement actions such as visits and phone calls without any penalty whatsoever. The costs of filing a civil case are substantial, and the criminal remedies are seldom used. From EPA's perspective, the important problem is not determining how high the penalties should be set, but deciding when to seek penalties at all. Observed penalty levels will depend very much on the resources available for such litigation.

The actual penalties assessed are often quite substantial, as we will illustrate with data from federal civil cases. (Penalties assessed in state courts are not tabulated by EPA.) Because of the lag between the initiation of civil litigation and the imposition of penalties, the prevailing penalty level serves as an index of the stringency of enforcement in earlier periods.

Penalties have been increasing, rising from $10 million annually in 1978 and 1980 to an average of $50 million annually in 1983 and 1984.[8] These increases reflect not only more stringent penalties for particular violations but also a broader scope of EPA operations and more vigorous enforcement efforts. From FY 1981 to FY 1985, the proportion of the EPA budget devoted to enforcement issues rose from 15 percent to 20 percent.[9]

To illustrate the range of penalties that can emerge for different degrees of violations of EPA standards, we will examine civil cases filed for water pollution violations in the pulp and paper industry, which includes approximately 350 major dischargers of pollutants.[10] Overall, there have been 40 closed civil cases against violators of EPA standards, with penalties assessed in 30 cases, ranging from $500 to $750,000. Eight firms were penalized by more than $100,000, and the average civil penalty levied was $115,000.

Thus EPA litigation efforts yield significant penalties for noncompliance. EPA enforcement policies can produce powerful incentives for compliance after violations have been discovered. But to the extent that litigation is pursued only for persistent violators of EPA regulations, the threat of a random initial inspection, however, is much lower. Moreover, the economic benefit to the firm from noncompliance may be substantial so that, on balance, even the fairly sizeable penalty amounts cited above may not provide effective incentives for compliance.

The ultimate test of the adequacy of EPA penalties is whether they induce compliance with government regulations. In this respect, the performance of EPA programs is quite mixed.

The compliance statistics cited below should be treated with some caution. Even if only a small percentage of firms are out of compliance, these firms tend to be the most significant polluters; consequentely any given percentage of noncompliance will be associated with a disproportionate amount of pollution. Moreover, the compliance data should not be treated as an unbiased measure.[11] Compliance judgments are typically based on unverified reports supplied by the polluting firms. For example, of the major stationary sources that EPA considers to be in compliance with its air pollution regulations, fewer than half have been audited in any way. Only for 5 percent of the firms has there been stack monitoring, and just under two-fifths have received an on-site inspection. Since many of these inspections are made with advance warning, available data very probably overstate the degree of compliance.

With these caveats in mind, consider first the degree of compliance with the state implementation plans for EPA air pollution regulations. For class Al sources of pollution (i.e., those with actual or potential emissions greater than 100 tons per year), there is 89 percent compliance with EPA standards.[12] Of the remaining 11 percent of the cases, about 4.2 percent of the firms have been found to be in violation of EPA standards, and for many the firm is believed to be out of compliance and a meeting with EPA has been scheduled.

Compliance with the water discharge permit program is comparable.[13] In this area firms that are not "significant" polluters are counted as being in compliance. The sample is restricted to major sources of pollution, that is, firms ranking among the lowest 15 percent in the national inventory of water dischargers. More than 90 percent of industrial polluters in this group are considered to be in compliance.

The rate of compliance of municipally operated facilities is much lower.[14] Only about 70 percent of municipal dischargers were in compliance, in the sense of having no significant violations. More than three-quarters (77 percent) of major publicly owned secondary treatment works had at least some noncompliance, and 53 percent had violations considered "serious." The greater prevalence of violations among municipally operated facilities may be due at least in part to EPA's general unwillingness to take enforcement actions against municipal governments.[15]

Compliance rates for more recently instituted EPA efforts are much lower. The compliance rates for hazardous and toxic substance regulations is believed to be only about 20 percent.[16] This low rate reflects both the rudimentary stage of the program's enforcement effort and the nature of the violation, which is often more clandestine and less readily observable than, for example, air pollution. In the case of asbestos removal in schools, where a stronger constituency is present to monitor compliance (parent and teacher groups), the compliance rates are greater (60 percent),[17] but still below those of the longer-established air and water efforts.

Overall, there is substantial, but not complete, compliance with EPA regulations. Although no study has formally shown the effectiveness of EPA's regulatory enforcement, from an economic standpoint one would expect the incentives created by EPA enforcement to influence the behavior of firms. Because EPA can assess significant penalties, major, continued violations of EPA standards are unlikely if EPA pursues legal action against the firm. Perhaps in part because of these sanctions, most firms are now believed to have met the minimum technological requirements imposed. Compliance problems persist, however, in the operation of pollution control equipment and fluctuations in its efficacy.[18]

II. The Ideal Policy Reference Point

A. The Basics of Optimal Regulation

From the standpoint of economic efficiency, EPA's objective should be to select policies that are in the best interests of society. More specifically, for concreteness we will take as the objective that the government should maximize the difference between the benefits and costs of the regulation, which is the net gain to society. This is a standard, though not uncontroversial, result in the policy analysis literature.[19] Here we will assume that various policy impacts have been stated in comparable terms, though in practice this may be far from straightforward.

Although the benefit-cost criterion has significant economic appeal and is convenient analytically, it is by no means the only approach that might be used. EPA could adopt a zero risk strategy, but doing so would be prohibitively costly. A less extreme alternative is to find an acceptable, non-zero level risk and impose that. Depending on the level selected, this strategy is not necessarily incompatible with a benefit-cost approach. Third, EPA could modify the benefit-cost framework to place greater weight on the losses to victims of environmental risks, In particular, it might weight the victims of involuntary risks differently from voluntary risks to take into account that compensation is not in practice paid to pollution victims. This absence of compensation is the greatest shortcoming of the standard benefit-cost approach, which implicitly assumes that compensation is paid by the generators of the environmental problems to those who suffer the damage. Finally, EPA might pursue a mixed strategy whereby it followed a benefit-cost approach possibly weighted by whether the risk was involuntary, subject to a requirement that the absolute level of risk not be too great. Here we will focus on the more standard benefit-cost criterion.

Figure 1 illustrates the incremental or marginal benefits (MB) from successive tightening of pollution control regulations as well as the associated marginal costs (MC). The marginal benefits are assumed to be a decreasing function of the level of pollution control because the initial reductions have the greatest impact on health and other environmental consequences. Similarly, the marginal costs of pollution control rise at an accelerating rate as the level of pollution control increases. The measures that produce pollution reductions most effectively are introduced first, and further reductions are increasingly costly. An additional cost component may be enforcement costs. These costs also tend to rise with the degree of pollution control since the most serious polluters can be identified most readily.

The optimal pollution level equates the marginal benefits and marginal costs of pollution control, so that the efficient level of pollution control is s*. This level could be achieved in one of two ways. First, the government could set a pollution control standard at s*. Firms would then be required to meet this level of pollution control, presumably by an enforcement system that requires compliance with the standard. Alternatively, the government could set a unit pollution tax of $v per unit of pollution and allow firms to pick their levels of pollution. In that case, the firm will equate $v to its marginal costs on a decentralized basis, and it will choose s* as the optimal level of pollution control.

What is noteworthy is the symmetry between the use of standards and penalties as a pollution control policy. This is the principal issue considered in the standards versus taxes literature. Although standards appear to be more rigid, they will produce the same environmental outcome as penalties. It should be noted, however, that this equivalence assumes that firms' decision regarding whether or not to remain in business is the same with standards and taxes. Since standards give the firm some free units of pollution, it may be that some inefficient firms remain in business because they are not paying their full costs of pollution. This global incentives issue is much less prominent in the literature, although it has been discussed by Carlton and Loury (1980).

Here our focus will be quite different. We will take as given EPA's decision to set environmental standards at some level s*. The question then becomes what penalty should be defined for deviations from the prescribed standard level. If EPA were to set a unit penalty at $v, it could be assured of achieving the optimal outcome, whatever the shape of MC. This is the lowest unit penalty amount that can achieve the optimal outcome.

B. Variations on the Basic Model

Heterogeneitv in Cost and Benefit Levels[20]

Setting uniform standards and penalties is appropriate if all firms have marginal costs curves such as MC and if the benefits of pollution control are identical across regions. Such uniformity is unlikely, however. Differences in the nature and vintage of technologies will make the marginal costs of pollution control quite different. Meeting a given water pollution standard may cost a pulp and paper mill much more than it does a corrugated box manufacturer. Rather than facing a single marginal cost curve, EPA may be confronted with a wide variety of such curves that in effect are averaged out across firms.

Differences in the benefits of pollution control also may arise because of differences in the size of the population exposed, the nature of the exposure (e.g., whether there are recreational facilities on the waterway), and the level of pollution from other sources.[21]

If there are differences in the health consequences of pollution from different firms, then the marginal benefit curve in Figure 1 also will differ. Higher marginal benefits imply that MB will shift upward,.leading to a more stringent standard s* or a higher penalty $v than would otherwise be desirable. Pollutants that affect more people or have especially severe consequences should be regulated more tightly. Procedures for accomplishing this objective for the case of benzene emissions are outlined by Nichols (1984).

EPA varies exposure limits for water pollutants, setting tighter restrictions on toxic wastes than on settleable solid wastes, for example. Broad geographic regions also have different air pollution control requirements. These differences did not arise because of differences in benefit levels, however, but because of a desire to prevent degradation of air quality in a region or perhaps to further some political objective, such as imposing lower costs on the industrial states of the Northeast and Midwest.[22] Overall, EPA makes no comprehensive policy effort to attune pollution control levels to benefit values.

Significant differences in the marginal costs of pollution control should also lead to differences in standards. Older facilities typically have higher costs of compliance because the pollution control devices must be integrated into an existing technology, which is usually a more expensive process than incorporating the necessary controls into the technology at the time of its design. Under a pollution charge system, older firms would choose higher levels of pollution than newer firms because of these differences. If we wish the EPA standards to replicate the outcome of a pollution fee system except for the difference in the effect of these policies on firms' closedown decisions, one should set tighter regulations for new sources of pollution than for existing sources.

EPA makes just this kind of distinction with respect to new source performance standards for air pollution. New sources are subject to much more stringent requirements than are existing sources. Major cases in point are the steel industry's total suspended particulate control requirements and the requirement that all new coal-fired utility boilers have desulfurization systems regardless of the sulfer content of the coal.[23]

Although such distinctions may have been politically rather than economically motivated, they are not necessarily undesirable. The direction of the much-bemoaned "new source bias" in EPA regulations is correct. The extent of the bias, however, may be excessive. The neglect of the benefits of regulation in setting the standard for coal-fired utility boilers, for example, suggests that the regulation may be too stringent, i.e., the incremental benefits of pollution control may be below the costs.

Uncertainty in Benefits and Costs

The benefits and costs associated with most EPA regulations are fairly uncertain. The overall levels of a pollutant often vary stochastically over time because of changing meteorological conditions or variations in the performance of pollution control equipment.[24] In addition, EPA's initial cost estimates are often highly uncertain, particularly when the proposed pollution control technology is not yet in use. The cost estimates that appear in analyses of regulatory proposals often convey an illusory degree of precision by failing to indicate the extent of our ignorance about likely compliance costs. For some risk regulations, the cost may be quite different from what was envisioned initially.[25]

Benefits may also be quite different. In some cases, the benefits depend on uncertain factors other than the regulation. This is particularly important when there is a no-risk threshhold, nonlinearities in the dose-response relationship, or the presence of multiple risks that are interactive rather than additive. The benefits from reducing asbestos exposures, for example, hinge on the level of cigarette smoking of those exposed to asbestos. In some instances our uncertainty is more fundamental. We may have some evidence that a particular substance is potentially carcinogenic, but have little idea of the dose-response relationship. In this case, we should use the mean of the distribution as our guide.

The literature on uncertainty with respect to benefits and costs sheds light on the relative performance of standards and penalty schemes in some extreme cases.[26] If the marginal cost curve for pollution control is relatively flat and constant, so that the principal uncertainty regards benefits, then standards will be more effective than pollution fees in ensuring the appropriate amount of pollution control. The reason is that any small variation in benefits, which are uncertain, will lead to a substantial variation in the optimal amount of pollution control if MC is flat. Similarly, a relatively flat marginal benefit curve coupled with highly uncertain marginal costs will make a penalty approach superior. Since neither extreme situation is likely to prevail in practice, the usual recommendation is to adopt a mixed system in the presence of uncertainty.

In general terms, the government now has such a mixed system. EPA enforces standards with penalties. Current policies differ from the economically efficient approach in two important respects, however. First, the standards are often set at levels much more stringent than s* in Figure 1 because of the nature of EPA's legislative mandate, as will be outlined below. Second, the penalties now levied are not designed to offer the polluter a choice as to whether to comply with the regulation. Rather, penalties are viewed as punitive measures to compel compliance, wherever the standard has been set.

The literature on regulatory uncertainty also offers some insight into appropriate penalty levels. Spence and Weitzman (1978) advocate a penalty level equal to the marginal benefits of pollution control at the current level of pollution, but Crandall (1983) correctly observes that such a penalty will be too high unless the penalty is successively lowered as the firm lowers its pollution level. Near the MB-MC equilibrium, however, the penalty based on the marginal benefits is a correct strategy. A penalty level of $v in Figure 1 is sufficient to provide incentives for efficient pollution control.

C. Incomplete Compliance[27]

In practice, policy outcomes are not achieved simply by issuing regulations. Instead, regulations specify guidelines for firms' actions, which must then be enforced by EPA. If enforcement is lax, there is no assurance that firms will comply with the regulation.

More generally, the normal approach is to assume that firms act in their self-interest. The high costs of compliance associated with very tight regulations and particularly expensive technologies are particularly likely to lead a firm to violate a regulation. Existing firms are thus less likely to comply with a new standard, particularly if they employ outmoded technologies. If the cost of the EPA regulation would undermine the viability of a marginal firm, the firm may decide to keep operating, in violation of the standard, until EPA seeks to bring it into compliance, and then to cease operations.

In such a situation of potential noncompliance, setting the level of a standard has a twofold impact. First, the stringency of the standard alters the optimal pollution level that firms complying with the standard will select. Second, the standard level will affect the likelihood that a firm will comply with the regulation. From the standpoint of environmental quality, tightening a standard consequently will have the effect of raising the level of pollution control for firms that choose to comply with the standard, and it will increase the likelihood that the firm will be out of compliance. As a result, standards do not necessarily lead to more pollution control. Moreover, from the standpoint of economic efficiency, tighter standards may be undesirable if they go below (i.e., to the right of) the pollution level s* in Figure 1.

In such a situation, penalties have a particularly important role to play. Higher penalties will make firms less likely to choose to violate a standard as it is tightened, thus reducing the loss in pollution control due to noncompliance.

D. Constraints on Efficient Policies

Institutional Constraints on Penalties

The most important institutional constraint on setting penalties is imposed by the statutory provisions governing civil penalties. Although the statutory penalty structures vary from program to program, we shall focus on two e-specially important ones -- those under the Federal Water Pollution Control Act (FWPCA) and under the Clean Air Act (CAA). For present purposes, two features of their penalty structures are especially relevant: (1) the respective roles of agency and court in assessing penalties, and (2) the substantive criteria that determine the penalty levels that may be assessed. It should be emphasized, however, that these statutory features apply to EPA's formal penalty-assessment processes. Because most enforcement is actually conducted informally rather than through formal adjudication, EPA's settlement practices, and not just its formal criteria, play an essential role in determining what penalties are actually imposed.

The Water Act[28]

Under the FWPCA, the EPA is not empowered to assess monetary penalties itself; it can only seek such penalties by bring in a civil action in court.[29] If EPA persuades the court that a violation has occurred, the court may impose a penalty on the violator of up to $10,000 per day.[30] (Alternatively, the agency may first merely issue an order requiring compliance; if that order is violated, the agency may then seek the civil penalty.)[31]

In deciding upon the penalty level under the FWPCA, the court enjoys complete discretion, for the statute prescribes no criteria for computing penalties. The court could impose the penalty proposed by the agency or impose a penalty designed to produce optimal levels of pollution, but it is not confined to these alternatives.[32] Some courts have ruled that the civil penalty should not be viewed as the functional equivalent of a damage remedy geared to the costs that the violator has imposed on others (although another court presumably remains free to take a contrary view). Rather, they have held, the penalty's purpose is to deter the offender and others from committing additional violations.[33] Some courts, however, have gone further and stated that the civil penalty should be used for punitive or exemplary purposes,[34] thereby justifying a penalty in excess of either the violator's benefits of noncompliance or the social costs of the prohibited pollution.

In setting penalties under the FWPCA, the court may -- but need not -- be guided by the EPA's "Civil Penalty Policy," which was issued in 1978 and is applicable *to* certain violators of the FWCPA and the CAA.[35] The Civil Penalty Policy enumerates four factors that are to be considered in computing the "minimum civil penalty" that would typically be presented to the court as the appropriate penalty (subject, of course, to the statutory ceiling of $10,000 per day). The four factors are: (1) "the harm or risk of harm to public health or the environment"; (2) the "economic benefit of delayed compliance" gained by the violator; (3) any "recalcitrance, defiance, or indifference" to legal requirements amounting to bad faith; and (4) any "extraordinary" enforcement costs borne by the government.

The Penalty Policy also recognizes two mitigating factors, which may be applied to reduce the recommended penalty level by that portion of the noncompliance attributable to the government itself, and by that portion attributable to other factors outside the violator's control. Finally, the minimum penalty may be further reduced in order to encourage settlement; in that event, the sum otherwise appropriate would be reduced by a factor reflecting the likelihood that the government will be unsuccessful in establishing the violation.

The EPA Penalty Policy is intended to affect three things: the amount of the penalties sought by EPA in civil actions, the amount acceptable to EPA in settlement of such actions, and the amount recommended to the court at trial. EPA conceives of the amounts determined under the policy not as effluent charges -- payment does not entitle the payor to continue polluting -- but only as deterrence designed to encourage compliance, (Of course, a punitive element is implicit in factor (3) above, relating to bad-faith noncompliance.) Again, however, the Penalty Policy is only advisory and precatory; it has no binding effect on the states, the courts, or even EPA itself. At least one state court has relied upon the Penalty Policy, but it did so only as a discretionary matter.[36]

Finally, Congress in 1977 expressly denied the EPA statutory authority under the FWPCA to impose administratively automatic "noncompliance fees" equivalent to the economic benefits of noncompliance. Althought the Senate added such a provision, it was deleted in conference as "unnecessary," in part (Senator Muskie said) because the EPA remained free to continue its enforcement strategy of seeking court-imposed penalties calibrated to the economic benefit of noncompliance, a strategy that Muskie said embodied congressional intent.[37]

In sum, the EPA does not control penalty-setting under the FWPCA. It does, however, influence penalty levels through its settlement actions and its proposals for court-imposed penalties, both of which are guided by an explicit, published Penalty Policy that places great (but not exclusive) weight on the economic criteria of pollution costs to society and noncompliance benefits to violators (at least up to the statutory ceiling of $10,000 per day). It would require new legislation to confer authority on the agency to assess penalties by itself, or to constrain the criteria employed by courts in setting penalty levels.

The Clean Air Act[38]

In contrast, the statutory penalty structure under the CAA is highly articulated and prescribes a central role for the EPA. With respect to most categories of CAA violations, the EPA is authorized by the CAA to bring a civil action to recover a court-imposed "noncompliance penalty" or, alternatively, to impose penalties administratively, i.e., without going to court.[39]

The CAA is silent concerning how a court is to determine the actual level of a civil penalty. But the statute does prescribe a criterion for the level of the noncompliance penalty: "no less than the economic value which a delay in compliance beyond July 1, 1979, may have for the owner of such source, including the quarterly equivalent of the capital costs of compliance and debt service over a normal amortization period, not to exceed ten years, operation and maintenance costs foregone as a result of noncompliance, and any additional economic value, which such a delay may have for the owner or operator of such source. . . ." From this sum must be subtracted any compliance expenditures already made by the owner or operator.[40]

This criterion has a number of noteworthy features. First, it is explicitly economic, defined by the noncompliance benefits gained by the owner or operator of the offending source. Second, this EPA-assessed penalty is not subject to any statutory ceiling on amount, while court-assessed penalties are confined by the CAA to a maximum of $25,000 per day.[41] Third, the CAA does not simply permit the EPA to assess noncompliance penalties; it appears to mandate their imposition. The statutory language provides that "the State or the Administrator shall, assess and collect a noncompliance penalty against every person who owns or operates" one of the enumerated stationary source categories.[42] Moreover, these stationary source categories are quite clearly and comprehensively defined; they are subject to only a few, narrowly-drawn exceptions.[43] Fourth, the amount of the penalty that the EPA is to assess is not discretionary; instead, it may be "no less than" the economic benefit of noncompliance, as defined in the statute and conforming regulations.[44] Finally, the penalty structure discourages dilatory legal challenges by polluters by providing that the penalties continue to accumulate while unsuccessful court appeals are pending.[45]

In 1980, the EPA issued regulations to implement the CAA's noncompliance penalty provisions.[46] In relevant part, the regulations require the owner or operator of a source that has received a notice of noncompliance from the EPA to calcuate the penalty in accordance with EPA instructions and criteria. The EPA's economic model, described in the Federal Register., provides these criteria. It takes into account two of the three statutory components of the noncompliance benefits to the polluter upon which the penalty is to be based -- the capital costs avoided, and the operation and maintenance costs avoided. The third statutory component -- "any additional economic value" resulting from the delayed compliance -- is excluded from the model as impossible to calculate.[47] The EPA regulations, and their underlying methodology, were challenged in the courts and were upheld by the U.S. Court of Appeals for the District of Columbia Circuit in 1983.[48]

Constraints on Standards and Their Enforcement

Setting penalties or standards to produce optimal levels of pollution control is not a policy option that EPA officials now have. Most important is that the legislative mandates for the different programs often define the nature of the policy choice process as one that is quite different from maximizing the difference between the benefits and costs.

EPA is often clearly prohibited from setting regulatory standards in a manner that equates the marginal benefits and marginal costs. In extreme instances, legislation may forbid EPA to consider costs at all. For example, in setting ambient air quality standards under the Clean Air Act, EPA is prohibited from taking cost considerations into account. EPA does, however, routinely take into account the "affordability" of its regulations, as there is a deliberate effort to avoid shutting down industries, except in extreme cases.

EPA has considerably more flexibility with respect to penalties. The upper limits on penalties are quite high and unlikely to be binding in most cases. There are no lower penalty limits.

These penalties could perhaps function as a pollution tax, to be paid if the firm chooses to pollute. Although many EPA officials regard penalties as a means for ensuring compliance, this view may stem primarily from agency tradition. In practice, EPA has many quite flexible penalty systems, but they are not identified as such.[49] The experimental marketable pollution rights scheme (or pollution reduction credits), which enables firms to trade pollution rights, in effect establishes a market for pollution, by imposing a pollution tax (i.e., the price of the credits) on all firms that purchase such credits. Although an overall pollution standard for which firms may purchase rights up to that overall standard preserves the image of a standards-based policy, the role of the standard is really not essential. As was shown above, standards and pollution fee systems simply approach the same problem in different manners.

Other related EPA policies also take a market approach to pollution. For example, firms can "bank" their emissions reduction credits, carrying over to future years the rights to pollution they were allowed in the past. The EPA "bubble" policy also offers some flexibility by allowing firms to trade air pollution rights internally, for example, by reducing emissions at one smokestack and increasing emissions from another.

As indicated above, some penalties are simply beyond EPA's control. For many civil and criminal sanctions, EPA must have the courts assess the penalties, where these cases are litigated by the U.S. Department of Justice. In these instances, EPA can recommend appropriate penalties, but it has less discretion than it would were the penalty system entirely under its administrative control.

III. Setting Penalties for Efficient Standards

A. Continual, Perfect Monitoring

In the simplest case for EPA penalty setting, the standards have been set at efficient levels (i.e., at the pertinent s* value in Figure 1) and EPA is always aware of firms' compliance status. The assumption of perfect monitoring is especially unrealistic because of the inherent difficulties of ascertaining whether the thousands of firms in each state are in compliance with the various EPA regulations. Clandestine disposal of toxic materials is particularly difficult to monitor.

In this section, we will begin by assuming the perfect monitoring case. Then we will analyze how the results change when monitoring difficulties are introduced.

If the EPA standards are set at the ideal level s*, the only function of the penalty system is to produce compliance with the standard. The unit penalty v* for any deviations from compliance is the lowest penalty that will ensure that firms take the needed actions. Any lower penalty will be exceeded by the marginal costs of pollution control for at least some level of pollution in excess of the allowable amount.

At the extreme, one could impose an arbitrarily large penalty on violations of the standard. But penalties larger than $v will have no influence on firms' choice of a pollution level. Punitive penalties that simply transfer resources from violating firms to EPA have no legitimate economic basis and are likely to meet with considerable resistance. Nevertheless, the government may wish to set a penalty in excess of the damage in order to punish the wrongdoers and establish effective incentives for compliance. Extreme penalty levels would, however, be perceived as unfair.

Even if EPA can monitor firms' actions perfectly and administer efficient penalties, firms will not always be in compliance. Compliance with a standard is a stochastic event that hinges on a variety of uncertain factors, such as the level of plant operations, equipment failure, and weather conditions. A penalty system severe enough to force universal continuous compliance would clearly be undesirable.

Problematic performance of pollution control equipment and other pollution-related factors can be diminished in terms of their importance, but often at a substantial cost. Clearly EPA should insist on a certain level of reliability in firms' pollution control. But as in setting the standard itself, EPA needs to consider the overall impact of greater reliability on the firm's likelihood of compliance. Greater pollution control benefits will result from increased compliance, but the costs of the greater reliability will also be greater. The ideal solution will strike a balance between these competing factors by allowing some occasional chance of noncompliance. We should encourage firms to take the measures needed for compliance, but we need not impose harsh penalties for a random inadequacy in the pollution control equipment.

B. Imperfect Monitoring

EPA has a variety of sources of information on firms' compliance status, including data supplied by firms and information gathered by the agency itself. In each case, however, there are substantial gaps in EPA's knowledge. In reporting their water pollution discharges, firms may report their minimum figures when averages are requested, or average amounts instead of maximum discharges. Even inspections by EPA are subject to error, depending on conditions on the particular day of the inspection and the thoroughness of the inspection.

Suppose there is some chance p that EPA will discover a particular violation. If firms are risk neutral, then EPA can mix penalties and imperfect monitoring to achieve the same results attainable under perfect monitoring by setting the penalty level at $v/p. For example, if $v is $5,OOO with perfect monitoring, it would be $10,000 for a p value of 0.5 and $50,000 for a p value of 0.1. In each case firms will face an expected penalty (i.e., probability of inspection multiplied by the size of the penalty) that is sufficient to lead them to choose s*.

Since it is costly to increase the probability p of identifying a violation, the ideal enforcement strategy is clear.[50] EPA should reduce its enforcement budget to an extremely low level and then raise penalties to an arbitrarily large amount. This solution would be ideal if firms were risk-neutral, but for extremely large penalties firms will be sensitive to the risk of a severe fine. The presence of risk aversion consequently will lower the optimal penalty level.

Fairness is a related concern. Government programs have a need for legitimacy, particularly since widespread perception that a regulation is being administered capriciously may undermine the credibility of the enforcement effort. The firm facing a large penalty could justifiably argue that if EPA’s enforcement effort were more effective it would face a lower penalty since the probability of detection would be greater, reducing $v/p. The firm should not be fully responsible for the low probability of inspection. A loss of credibility from inappropriate penalties may influence reform of the agency's enabling legislation, its budget, and the degree to which public knowledge of a violation itself serves as a sanction.

Presumably, as in many judicial contexts, the size of the penalty should bear some relation to the amount of harm done. A firm causing $5,000 in environmental damages clearly should not be penalized $5 million, no matter how small EPA's likelihood of discovering the violation.

Two kinds of legal constraints also bear on penalty levels. First, certain limits are imposed by the enabling legislation for specific EPA programs. Second, a firm can reorganize or declare bankruptcy to limit its losses if the penalties become particularly severe. This problem has received most attention in connection with the product liability claims for asbestos victims;[51] it could arise with respect to particularly severe regulatory penalties as well.

C. Heterogeneity in Penalty Levels

Under ideal conditions an EPA standard should vary with marginal benefit and marginal cost levels at a particular firm. The unit penalty level $v needed to ensure efficient compliance would also vary.

Because the optimal penalty depends on the marginal benefits at the optimal level of pollution, one would expect the amount of the penalty to increase with the size of the exposed population, the degree of exposure, or the severity of the health outcome. Similarly, it is efficient to relax standards when marginal costs of compliance are high. If one is trying to replicate with penalties the outcome under a standard based on the firm's specific marginal benefits and marginal costs, then the optimal unit penalty should also be reduced when marginal costs are high, since the appropriate pollution control objective in the presence of high compliance costs will be less. EPA consequently should set lower penalties for older firms or for firms where compliance involves a major transformation of the technology rather than, for example, an end-of-pipe water pollution control measure.

IV. Optimal Penalties with Inefficient Standards

In the usual case, EPA bases penalties on risk minimization concerns rather than on a comparison of benefits and costs. As a result, the standards often may be too tight based solely on economic efficiency grounds. Although EPA may wish to enforce such standards at the levels at which they have been set, it may be that the agency has a broader policy objective. The standards may have been set in an earlier period when there was less concern for the need for balanced policies because the limits on our resources for reducing risk were not well understood. Alternatively, EPA policymakers may be fully cognizant that the standard is too stringent from the standpoint of economic efficiency, but they may nevertheless wish to enforce such stringent regulations because of other policy objectives, such as a desire to prevent people from being exposed to major risks. Overly stringent regulations also can be enforced selectively by imposing substantial penalties for large violators and perhaps ignoring small violations altogether. Here we will outline possible approaches to-setting penalties that depend on one's particular policy objective.

Figure 2 illustrates the case of inefficiently stringent standards: s^{**} rather than s^*. What penalty should EPA set for violations of this standard? Here we will focus on unit penalties based on the extent of the deviation from the value of s^{**}. In addition, we will ignore problems of the probabilistic monitoring of compliance status.

With efficient standards, penalties are keyed to the marginal benefit of pollution reduction. Firms then have an incentive to balance their costs against the benefits to society, thus internalizing the benefit-cost tradeoff on a decentralized basis. Once EPA sets standards that diverge from this efficient level, the conceptual foundation for setting penalties becomes distorted.

The five key points that will be discussed are marked A through E on the diagram. In each case, the discussion will pertain to the value along the vertical axis for each point, which represents the dollar value of either the marginal benefit curve or the marginal cost curve at the point.

The firm's marginal costs of additional pollution removal are given at point E in Figure 2. Any unit penalty amount above this level will lead the firm to augment its pollution reduction effort. Suppose the penalty is set at the level of point C, as in the efficient standards case. The firm will invest in pollution reduction up to s*, and it will pay a unit penalty based on C for the discrepancy between the s** and s* levels of pollution control. In effect, EPA will be transforming an inefficiently stringent standard into an efficient one, with the only difference being that EPA collects a pollution tax for pollution reductions below s**.

Although a firm at point C is not in compliance with the standard, EPA need not classify such deviations as significant noncompliance and need take no enforcement action beyond continued penalties. The penalties collected could be used to support more cost-effective government policies.

Establishing a unit penalty level only for deviations from a standard s** has the additional benefit of being less costly to a firm than a simple pollution tax[52] based on the firm's entire pollution. In effect, firms get the pollution levels between s** and full cleanup for nothing.

Firms that remove pollution up to level s* will reach the nonsignificant violator category and pay only nonpunitive penalties. What should be done about recalcitrant firms that fail to meet this pollution control level? Their own economic self-interest should lead them to make the needed investment if any penalty C or greater is imposed. If pollution reduction up to s* does not occur, it may be that the firm's marginal cost of compliance has been underestimated.

Other penalty levels based on the marginal benefits of pollution control have little potential role to play. Penalties set at A, which is the marginal benefit of pollution control at current pollution level, will lead to pollution cleanup beyond s* and, depending on the shape of the MB and MC curves, may impose a unit penalty higher than needed to ensure complete compliance at s*, since A may lie above B. Similarly, the marginal benefit value at s*, which is given by point D, will not even lead to efficient degrees of pollution reduction at C. Indeed, D may lie below E, in which case firms would not alter their level of pollution control at all in response to the penalties.

If the standard itself is not set at the efficient policy level, basing the penalty level on pollution reduction benefits is not a particularly useful approach to enforcement. Incorporating benefit concerns into the penalty setting process will create a different emphasis from that of the original standard, which ignored benefit-cost tradeoffs.

Unit penalties for noncompliance linked to point B will be sufficient to induce compliance at s**. Since the MC curve lies above MB for all points to the right of point C, no benefit-based measure could lead to a pollution level such as B.

While B represents the lowest unit penalty that will induce compliance, there is a cheaper penalty structure that will have the same result. If EPA penalizes firms for the cost savings from noncompliance, which is the area under ECB, then the economic incentive for noncompliance will be removed and firms will pick s**. The total cost to the firm will be less than if the firm had paid a unit penalty B for its violation amount s** - s*.

The cost savings to the firm, or what EPA terms the benefits of noncompliance, is a prominent measure used in EPA penalty setting.[53] This approach is in many respects well grounded if one's objective is to attain compliance at s**, a level beyond the efficient degree of control.

The EPA benefits-of-noncompliance approach has been defended on the ground that if benefits and costs are closely correlated, the EPA policy may produce an outcome similar to that of an ideal pollution tax.[54] Such a rationalization of EPA's policies misses their fundamental difference from an efficient penalty approach. The marginal incentives for pollution control under the EPA benefits-of-noncompliance approach will lead to pollution control up to point B, whereas efficient penalties will lead firms to move only to C. Whether the total penalty for noncompliance is higher or lower for the cost-based or efficient penalty approach is not the major issue. A firm at s* may pay more or less penalty under an efficient unit penalty than under a cost-based penalty. The fundamental issue is how the incentives are structured for improvements in pollution control beyond s* and for the level of pollution control ultimately selected.

If the policy intent is truly to promote pollution control up to s**, then the cost gains from noncompliance are a sound basis for policy. But such an outcome is not a good approximation of what would happen if penalties were set at efficient levels. That outcome is quite different and can be achieved by penalties set according to the marginal benefits of noncompliance at the efficient standard level s*. Coupling an overly stringent standard s** with efficient penalties is an attractive policy for promoting efficient enforcement, since the effective pollution tax is less than if firms had to pay for all their pollution, rather than only that above the standard.

It is- also possible to set a standard that is not stringent enough. In that case, little can be accomplished through a penalty system for noncompliance. Once a firm meets the standard, it presumably cannot be penalized. The standard itself would have to be changed to induce the firm to move to an efficient level of pollution control.

Because the penalty policy can remedy the shortcomings of overly stringent standards but not overly loose standards, one might envision the following approach to policy design. Optimal environmental policies are uncertain and are evolving over time. EPA can hedge against these risks by adopting very tight standards and adjusting the penalty level for violations to accommodate changing policy needs. Such revisions should not be erratic, however, because of the costs that uncertain regulatory policies impose on firms.[55]

Such a policy mix would be appropriate if, as seems likely, there is greater flexibility in setting penalties than in revising standards. Standards cannot be altered without a lengthy rulemaking process, whereas penalty levels are set administratively or by the courts. Similarly, permit levels for particular firms tend to become property rights that are difficult to reduce. In practice, however, both penalties and standards are likely to be difficult to vary over time because of political and administrative constraints. The agency's greatest leeway is with respect to both the strictness of inspections and prosecution and the targeting of these efforts. Despite their rigidity, standards may offer substantial flexibility through manipulation of the level and direction of the enforcement effort.

V. Operational Problems

A. Information Problems

Basing EPA penalties on either the costs or the benefits of noncompliance requires at least some information on what these values are. Yet these amounts may vary by firm and may not be readily attainable by EPA.

In the case of benefits-based penalties, EPA must first ascertain the amount of the firm's pollution and its difference from the standard. Water pollution amounts must be reported regularly to EPA in discharge monitoring reports. The level of the discharge can then be compared with the permit level to ascertain the extent of noncompliance. At the other extreme, firms' disposal of toxic chemicals is seldom reported and can usually be identified only for particular cases of extreme violations.

Once the extent of the violation is known, EPA must ascertain its benefit value, which typically means society's willingness to pay for the benefit. Because environmental amenities are not bought and sold in the marketplace, economists have sought more ingenious means for ascertaining such values. One approach is to analyze implicit valuations in market transactions. People do not purchase air pollution reductions, but they do buy houses whose property values are affected by the degree of pollution. Similarly, we do not purchase life extension reductions, but workers do receive additional wages for job risks, thus establishing the rate of tradeoff.

If no convenient market reference point exists, one can simulate a market through contingent valuations, that is, by asking people how much they would pay for certain attributes if such a market did exist.[56] In this way one can attempt to value attributes ranging from greater visibility at national parks to the value of preventing a child poisoning from a pesticide. The recent spate of research on benefits sponsored in large part by the EPA Benefits Office suggests that such issues are now considered more tractable than they were a decade ago.

Cost-based measures for assessing penalties do not eliminate all of the difficulties associated with selecting the penalty level, but shift attention to a new class of concerns. Costs of compliance are perhaps easiest to assess if the firm needs only to purchase some well-specified add-on technology, such as end-of-pipe treatment equipment. Although computation of the penalty is complicated by investment tax credits, the role of depreciation, and similar factors, the task is amenable to fairly precise analysis.

In the usual cases, however, the required improvements demand some modifications within the firm's technology. Operating conditions may also have to change if the firm is to meet the standard. Assessment of the costs of compliance in such instances may require detailed firm-specific knowledge. One possible solution is to adopt a truth-inducing mechanism whereby firms must report these costs subject to a penalty if the actual costs deviate too much from those that are stated. For a benefits-based penalty, the major uncertainty is not firm-specific but pertains to general issues such as the value of preventing a case of cancer.

Which policy approach involves more informational difficulties is not at all clear-cut. More fundamentally, it makes little sense to adopt an inferior approach to setting penalties for pragmatic reasons. In effect, one would be measuring an incorrect penalty measure precisely rather than the correct penalty value less precisely. In addition, the extent of EPA's knowledge depends in part on the policy approach that is adopted, since this will affect the focus of the field stff's efforts and the research funded by EPA.

Though it is important to recognize the information problems that will be encountered, the choice of EPA's policy approach to penalties should turn on more fundamental concerns, such as those raised in Section 4.

B. Legality and Legitimacy

The penalty structure must be legally feasible and must be generally viewed as a legitimate approach. Neither consideration will determine the choice between penalty policies discussed here.

Legally, EPA has upper limits on the penalties it can set, but it is by no means required to set penalties to ensure full compliance. The agency already distinguishes between significant and insignificant violations. Moreover, it has adopted a variety of innovative regulatory approaches, such as tradeable pollution permits, that are in the spirit of a pollution fee approach. There are no legal barriers to pollution fees, although the EPA tradition has leaned toward punitive fines to ensure compliance.

Neither penalties designed to ensure efficient pollution control nor those designed to ensure compliance can be rejected on legitimacy grounds. In the case of pollution fees, the penalty is linked to the harm done to society, a concept with a long history in the American legal system. Penalties based on the firm's gains from noncompliance will also seem legitimate in a society that wishes to prevent firms from profiting from violations of its regulations.

Overall, the choice between penalties for efficient pollution and penalties to ensure compliance will have little effect on the legality or legitimacy of EPA actions. The principal issue at stake is what emphasis EPA wishes its policy to have.

VI. Conclusion

In setting the penalty level, two fundamental approaches are possible. First, one could set the penalty so as to remove the economic gain to the firm from noncompliance. The manner in which EPA now sets penalties to accomplish this objective is a relatively attractive approach if the objective is to promote compliance. An alternative is to impose benefits-based penalties for deviations from a standard. Such a policy can induce efficient degrees of compliance and will be less costly than a standard polution fee system, since the firm pays only for pollution in excess of the standard. Under such a system, coupling penalties with "overly stringent" standards is an attractive policy mix. Which policy approach is preferable depends on whether EPA's objective is to ensure full compliance or to promote efficient degrees of compliance that recognize the tradeoffs involved. In either case, the penalty structure alone can be a significant instrument for achieving EPA's policy objectives.

Footnotes

[1]Federal Register, Vol. 49, No. 193 (1984), p. 39064.

[2]Wasserman (1984), p. IV-14.

[3]Ibid., pp. VI-18 to VI-19.

[4]Ibid., p. IV-20.

[5]Ibid;, p. IV-20.

[6]Ibid., p. IV-22.

[7]Ibid,, p. IV-20.

[8]Ibid., p. VI-24.

[9]Ibid., p. VI-25.

[10]These data are based on unpublished EPA computer printouts generated from the case file for this study.

[11]Crandall (1983), p. 29.

[12]These data, from Wasserman (1984), p. VI-5, are for 1983, but the results for 1984 are similar except that detailed breakdowns for the composition of the noncomplying set of firms are not available.

[13]The data below are based on Wasserman (1984), p. VI-6.

[14]Ibid., p. VI-7.

[15]Ibid,

[16]Ibid., p. VI-9.

[17]Ibid,

[18]Russell, Harrington, and Vaughan (1985).

[19]See Stokey and Zeckhauser (1978) for a general treatment, Baumol and Oates (1979) for an environmental policy discussion, and Chapter 7 of Viscusi (1983) for a series of examples from OSHA policy.

[20]The development in this subsection is based in part on Viscusi and

Zeckhauser (1979).

[21] In some cases, the heterogeneity in the benefits of pollution reduction can be eliminated from the analysis through appropriate scaling of the pollution control metric. Rather than have the horizontal axis in Figure 1 be the amount of pollution control, one could use as th scale the number of cases of cancer prevented or some other appropriate health outcome. In that instance, there is more likely to be much greater uniformity across firms' benefits of pollution control. Such a reformulation also recasts the nature of the penalty-setting process so that there is no real gain from redefining the problem in this manner.

[22] Robert Crandall (1983).

[23] Ibid,, pp. 39-44.

[24] See Russell, Harrington, and Vaughan (1985) and Watson and Downing (1976) for discussion of the variations in the efficacy of pollution control.

[25] See the discussion in Viscusi (1985) of the substantial overstatement of the costs of the OSHA cotton dust standard.

[26] See Weitzman (1974), Roberts and Spence (1976), Spence and Weitzman (1978), and Crandall (1983).

[27] This section is based on the results in Viscusi and Zeckhauser (1979).

[28] 33 U.S.C. Section 1251 et seq,

[29] 33 U.S.C. Section 1319(b).

[30] 33 U.S.C. Section 1319(d).

[31] 33 U.S.C. Section 1319(a)(3).

[32] E.g., U.S. v. Velsicol Chemical Corp., 8 Envtl. L. Rep. (Envtl. L,. Inst.) 20745, 20748(W.D. Tenn. Aug. 31, 1978); U.S. v. Cargill, 508 F. Supp. 734, 751 n. 59 (D.Del. 1981).

[33] E.g., U.S. v. Lambert, 13 Envtl. L. Rep. (Envtl. L. Inst.), 20489 (M.D. Fla. 1983); U.S. v. Velsicol, Supra.

[34] E.g., U.S. v. Velsicol, supra; State ex rel. Brown v. Dayton Malleable, Inc., 438 N.E. 2d 120, 1 Ohio St. 3rd 151 (Ohio 1982).

[35] 45 Fed. Reg. 50110 (1980), codified at 40 C.F.R. Part 66 (1980).

[36] State ex rel. Brown v. Dayton Malleable, Inc., supra.

[37] See 1977 Legislative History, vol. 3, at 476-7 (1978).

[38] 42 U.S.C. Section 7401 et seq. (1977).

[39] 42 U.S.C. Section 7420(a)(2) (1977).

[40] 42 U.S.C. Section 7420(d)(2) (1977):

[41] 42 U.S.C. Section 7413(b) (1977).

[42] 42 U.S.C. Section 7420(a)(2)(A) (1977).

[43] 42 U.S.C. Section 7420(a)(2)(B) and (C) (1977).

[44] 42 U.S.C. Section 7420(d)(2)(A) (1977).

[45] 42 U.S.C. Section 7608(g) (1970).

[46] 45 Fed. Reg. 50110 (1980) codified at 40 C.F.R. 66.1 et seq. (1980).

[47] See 45 Fed.Reg. 50086, 50090 (1980).

[48] Duquesne Light Co. v. EPA, 13 E.L.R. 20251 (D.C. Cir. 1983).

[49] For a review see Crandall (1983).

[50] See Polinsky and Shavell (1979) for treatment of the issues in this paragraph.

[51] See Viscusi (1984).

[52] A similar concern with the costs to firms is a principal factor underlying Crandall's (1983) analogous proposal that a two-part tariff be used for air pollution.

[53] Susan Dudley, "Economic Benefit of Delayed Compliance," internal EPA memorandum, September 1984.

[54] See Drayton (1980).

[55] See Viscusi (1983).

[56] W. H. Desvouges, V. K. Smith, and M. McGivney, "A Comparison of Alternative Approaches for Estimating Recreation and Related Benefits of Water Quality," Environmental Benefits Analysis Series, U.S. Environmental Protection Agency, Washington, D.C., 1983.

References

Baumol, William, and Wallace Oates, Economics. Environmental Policy. and the Quality of Life (Englewood Cliffs: Prentice-Hall, 1979).

Crandall, Robert, Controlline Industrial Pollution: The Economics and Politics of Clean Air (Washington: Brookings Institution, 1983).

Carlton, Dennis, and Glenn Loury, "The Limitations of Pigouvian Taxes as a Long-Run remedy for Externalities," Quarterly Journal of Economics, Vol. XCV, No. 3 (1980), pp. 559-566.

Desvouges, W. H., V. K. Smith, and M. McGivney, "A Comparison of Alternative Approaches to Estimating Recreation and Related Benefits of Water Quality," Environmental Benefits Analysis Series, U.S. Environmental Protection Agency (1983).

Drayton, William, "Economic Law Enforcement," Harvard Environmental Law Review, Vol. 4 (1980).

Polinsky, A. Mitchell, and Steven Shavell, "The Optimal Tradeoff Between the Probability and Magnitude of Fines," American Economic Review, Vol. 69 (1979).

Roberts, Marc, and Michael Spence, "Effluent Charges and Licenses under Uncertainty," Journal of Public Economics, Vol. 5 (1976).

Russell, Clifford, Winston Harrington, and William Vaughan, Monitoring and Enforcement in Pollution Control, draft book manuscript, Resources for the Future (1985).

Spence, Michael, and Martin Weitzman, "Regulatory Strategies for Pollution Control," in Ann Friedlaender, ed., Approaches to Controlling Pollution (Cambridge: MIT Press, 1978).

Stokey, Edith, and Richard Zeckhauser, A Primer for Policy Analysis (New York: W. N. Norton, 1978).

Viscusi, W. Kip, "Frameworks for Analyzing the Effects of Risk and Environmental Regulations on Productivity," American Economic Review, Vol. 73, No. 4 (1983).

_____, Risk by Choice: Regulating Health and Safety in the Workplace (Cambridge: Harvard University Press,, 1983).

_____, and Richard Zeckhauser, "Optimal Standards with Incomplete Enforcement," Public Policy, Vol. 27 (1979).

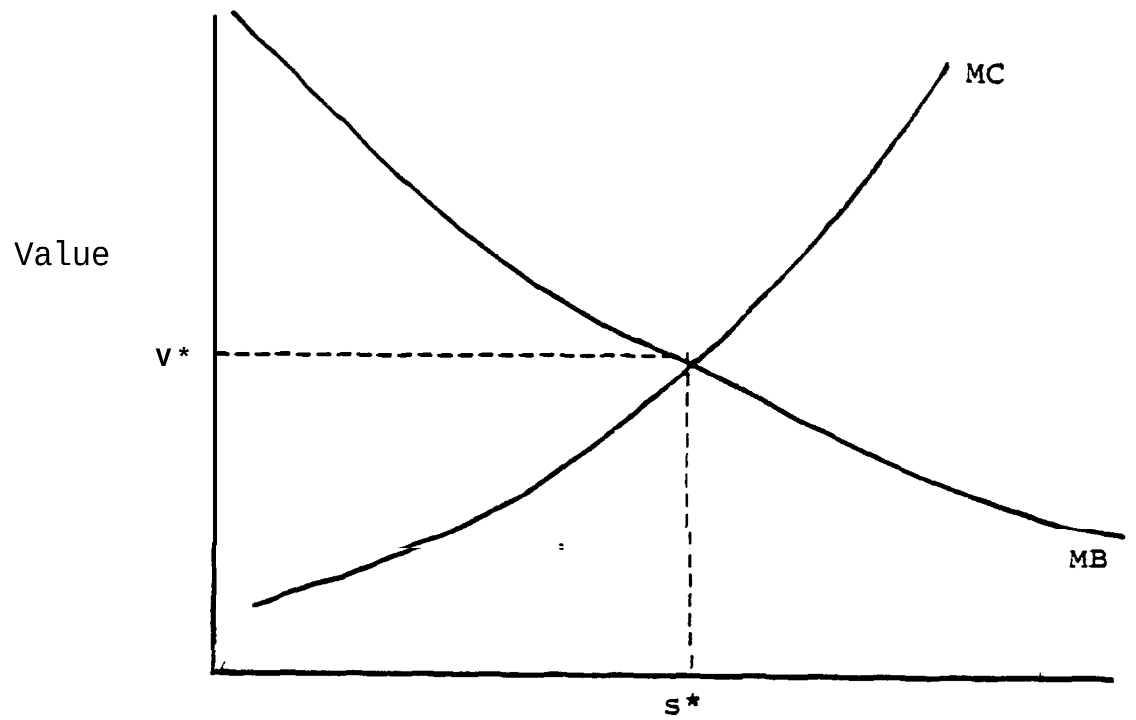
Wasserman, Cheryl, "Improving the Efficiency and Effectiveness of Compliance Monitoring and Enforcement of Environmental Policies," U.S. Environmental Protection Agency (1984).

Watson, William, and Paul Downing, "Enforcement of Environmental Standards and the Central Limit Theorem," Journal of the American Statistical Association, Vol. 71 (1976).

Weitzman, Martin, "Price vs. Quantities," Review of Economic Studies (1974).

Figure 1

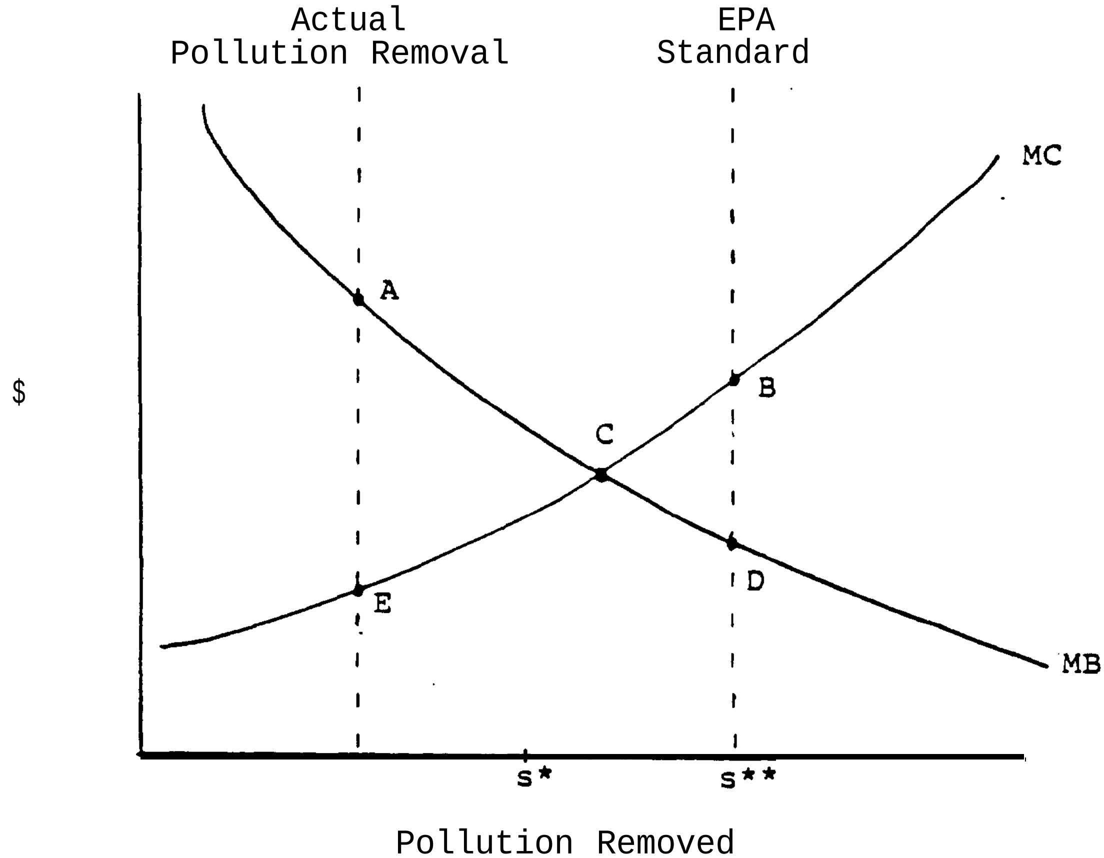
Setting Optimal Regulations



Amount of Pollution Control

Figure 2

Penalties with Overly Stringent Standards



TABL 1

SUMMARY OF PENALTIES AND IMPRISONMENT PROVISIONS

(In Order of Magnitude)

Statute	Civil		Criminal		
	Citation	Penalty (Not more than)	Citation	Penalty (Not more than)	Imprisonment (Not to exceed)
1. Safe Drinking Water Act	1414(b)	$5,000/day			
	1423(b)	$5,000/day	1414(c)	$5,000/day	None
			1423(b)(2)	$10,000/day	None
2. Superfund	109	$10,000/day	103(c)	$10,000/year	1 year
			103(d)	$20,000/year	1 year
3. Clean Water Act	309(e)	$10,000/day	309(c)(1)	$2,500-$25,000/day	1 year
				Second violation: $50,000/day	2 years
			309(c)(2)	$10,000/day	6 months
4. Toxic Substances Control Act	16(a)	$25,000/day	16(b)	$25,000/day	1 year
5. Clean Air Act	167	$25,000/day	113(c)(1)	$25,000/day	1 year
	205	$10,000/day		Second violation: $50,000/day	2 years
	303(b)	$5,000/day	113(c)(2)	$10,000/day	6 months
6. Resource Conservation and Recovery Act	3008(a)(3)	$25,000/day/violation	3008(d)	$50,000/day	1 year
			3000(d)	$25,000/day	2 years
	3013(e)	$5,000/day	3008(e)	$250,000	2 years
			3008(e)	$250,000	5 years
			3008(e)	$1 million	

Source: Cheryl Wasserman, "Improving the Efficiency and Effectiveness of Compliance Monitoring and Enforcement of Environmental Policies," U.S. Environmental Protection Agency, 1984.

Chapter 4

THE APPROPRIATE REGULATION 0F INVOLUNTARY AND HIDDEN RISKS*

by

W. Kip Viscual

University of Chicago (Visiting)

and Northwestern University

and

Richard J. Zeckhauser

Harvard University

*This research was supported under an EPA Cooperative Agreement from the Office of Policy Analysis. Viscusi’s research was also supported in part by the Center for the Study of the Economy and the State, University of Chicago.

I. Introduction

Risks differ not only in terms of their magnitude and their welfare implications, but also in terms of the process by which individuals incur the risk. At one extreme, the decision to bear the risk may be the consequence of an entirely voluntary market tranaction. A worker who is fully cognizant of job risks will receive additional wage premiums that restore his alternative utility level, and new purchasers of homes located near airports will require a price differential to compensate for the additional noise pollution. In the case of job risks the party generating the risk is involved in the transaction, creating market incentives for safety, whereas in the airport example the airport does not bear the cost of the externalities generated. In each case, the individuals bearing the risk are compensated, although the original homeowners who were present before the airport was built will suffer a loss from the decline in property values.[1] Once the individual has moved near the airport and been compensated, his welfare will be reduced by each successive increment in the noise level, however. Moreover, the market serves a constructive function in leading those less sensitive to noise to sort themselves into living near the airport.

In contrast, for risks that are entirely involuntary there is no explicit or implicit compensation for the risk. Indeed, it is this absence of the market that is a driving force behind government intervention based on market failures. In this paper we will explore various consequences of the absence of market transactions for involuntary risks as well as the specific characteristics of the regulation of these risks.

The voluntary/involuntary risk dichotomy will greatly facilitate the analysis, but it should be understood that there is a continuum of such cases.[2] Individual knowledge pertaining to market traded risks may be imperfect, possibly impeding the efficacy of market transactions, but nevertheless being sufficiently correlated with the true risk that the market plays a constructive role. Similarly, there may be transaction Costs to individual mobility that limit the degree of effective choice that individuals may have. For instances that fall between the extreme situations of completely voluntary and totally involuntary risks the discussion below can be modified appropriately.

The only truly voluntary risks are those for which there is no capital component. With jobs there is an involuntary aspect to the extent that there is a cost of switching jobs. If jobs worked as a daily spot market those concerns would be less prominent than if seniority and other impedimenta to mobility play a prominent role.

That involuntary risks are more likely to be candidates for government intervention is not a surprising observation. What we will consider is the related issue of how stringently these risks should be regulated, given that the government has decided to regulate the particular risk. In terms of the target risk level that it is optimal for the government to enforce through its regulations, should a distinction be made between voluntary and involuntary risks? This distinction is much different from whether the amount of risk reduction should be greater for involuntary risks. The absence of a market makes the differences between the observed risk level and the target risk level after regulation greater for involuntary risks, so that even if the target risk level is unaffected the amount of risk reduction will differ. What we will suggest is that for a diverse set of reasons based on both efficiency and equity grounds not only will the optimal degree of risk reduction for involuntary risks be greater than for an equal voluntary risk, but the optimal post regulation risk level will be lower as well.

There are three classes of reasons why we should differentiate the treatment of voluntary and involuntary risks: the mix of individuals bearing the risk differs, equity considerations may be of consequence, and there may be inadequacies in both the structure and the enforcement of regulatory Policies. What is most striking is that all of the various influences cited suggest that more stringent regulation of involuntary risks is warranted.

II. Differences in the Valuations of Risk Reductions

Heterogeneity of Risk Preferences[3]

Following the standard principles of policy analysis, the value of reducing a particular risk is society’s willingness to pay for the risk reduction of all members of society. A chief component of this measure will be the willingness to pay of the individual directly affected by the risk. In general this value will differ between involuntary and voluntary risks because the group of individuals choosing to incur risks on a voluntary basis will have different risk dollar trade offs.

In particular, in the case of voluntary risks one would expect those with comparatively low values of life to self select themselves into the higher risk pursuits. Empirical results for fatal job risks suggest that workers in high risk jobs with annual death risks on the order of 1 in 1000 have implicit values of life reflected in their wage risk tradeoffs of $600,000.[4] In contrast, workers in typical blue collar jobs with a risk of 1 in 10,000 have a value of life of averaging $3.5 million, and workers in safe jobs have a value of life on the order of $6lO million.

The diversity of these estimates suggests that there may be quite substantial differences in the willingness to pay for the reduction of voluntary and involuntary risks. If the voluntary risks are nontrivial, the nature of the self selection process generating the mix of persons bearing the risk implies that they will tend to be borne by individuals on the low end of the value of life spectrum. Involuntary risks are spread across the entire population so that the appropriate value of life will reflect a weighted average of these individual valuations, which will generally be higher than the average valuation of individuals who have knowingly decided to accept substantial risks.

If individuals' values of life were uniformly distributed across the population, the value of life of the median individual would be the appropriate benefits measure. Empirical evidence suggests, however, that the distribution of values of life consists of a bottom quartile of the Population with low values of life and the remaining three&fourths of the population who have similar values of life but which are higher than those in the bottom quartile. As a result, for involuntary risks the value of life of the median individual will tend to overstate the average value of life for the population affected by an involuntary risk. This amount in turn will be greater than the value that would pertain to voluntary risks. Because of the nature of the job risk data on which these calculations are based, there may be even more heterogeneity among the affluent segment of the population than has previously been estimated.

Recognition of variations in willingness to pay for risk reduction sidesteps an important income redistribution issue. It may, for example, be efficient to reduce air pollution in affluent areas but not in poor areas because of differences in benefit valuation. This procedure would be clearly desirable if in fact cash compensation were paid to the poor in lieu of greater risk reduction. Targeting policies on the basis of differences in benefit valuations across income classes is only compelling if income redistribution efforts or the progressivity of the tax system is designed to take into account this differential emphasis.

Difference in Attitudes toward Involuntary Risks

Although the usual assumption in the risk literature is that individuals have given risk preferences that are unaffected by whether or not there is a market transaction leading to the imposition of the risk voluntarily, attitudes toward risk bearing may be quite different depending on whether there is such a transaction leading to explicit or implicit compensation for the risk.

The individual bearing a given risk may be willing to do so for some additional compensation, but if he does not receive this compensation his Welfare loss may be more than this amount. More specifically, the importance of the process by which economic outcomes are generated may have a significant bearing on the associated welfare implications of a particular risk. Because there is no compensation for involuntary risks, the absence Of a process that maintains the individual's welfare level through a market transaction may itself be viewed as an attribute of the risk that enters the individual's utility function. We live in a society that believes that exchange should be voluntary, and this principle pertains to risks as well.

The degree of volition and compensation for the risk is likely to be a Particularly pertinent concern influencing general externalities to society at large. One would expect that society's altruistic concerns with respect to individuals affected by a hazard to be much stronger if the risk is being imposed voluntarily. In contrast, if workers are receiving substantial hazard pay that is sufficient to maintain their level of utility at the no risk level, then they are suffering no welfare loss. Altruism regarding persons bearing risks that impose no welfare loss on the victim because of the presence of risk compensation may be viewed entirely differently than involuntary risks for which there is no compensation and an expected welfare loss.

In terms of empirical predictions, one would expect public support for efforts to make cigarette smoking safer (by, for example, discouraging the smoking of cigarettes or restricting their use) to be less than the efforts to promote food safety. Such distinctions are evident in the emphasis of government policies.

As a result, the risk dollar tradeoffs pertaining to involuntary risk should be higher for two general classes of reasons. First, the mix of individuals affected by the risk will tend to have higher risk valuations on average. Second, both for the individual bearing the risk and society's general concerns with respect to the absence of a market transaction and the payment of compensation leads to a higher valuation of involuntary risks.

This latter factor can be viewed as concern with the security of the contract for the one area we may well value the most human health.

III. Equity Concerns

Conventional Equity

A Possible objective of any government program in a second best world of inadequate cash transfers is to promote income redistribution. If risks hurt the poor, lowering their welfare, we may choose to regulate these risks more stringently as a form of in kind redistribution.

In the case of involuntary risks, the potential role for risk regulation as a form of redistribution is relatively clear. More stringent regulations will reduce the risk that is imposed, enhancing the expected health of the affected individuals. Because there is no compensation for the risk, the beneficiary's welfare will be unambiguously enhanced by the regulation.

The desirability of risk regulation as a redistributive mechanism for voluntary risks is much less pronounced. The reduction of the risk leads to an expected improvement in health, as with voluntary risks, but it will also eliminate the market compensation for the risk. As a result, the financial resources of the affected parties will be reduced, which will always diminish the usefulness of risk regulation as a redistributive mechanism.

Indeed, from the standpoint of the individuals accepting a risk voluntarily, their expected welfare levels will be reduced by the presence of a risk regulation. From this standpoint, the regulation is eliminating the possibility of a market transaction that they found attractive. At the very minimum they were indifferent to bearing the risk and for those earning economic rents from the voluntary risk there will be a perceived welfare loss from the risk regulation.

It may be that the individuals' perceptions of the risk are erroneous, so that using the true probabilities a regulation may boost their actual expected welfare level. Nevertheless, the reduction in compensation for a risk incurred voluntarily will tend to make regulation of voluntary risks less desirable than comparable regulation of involuntary risks.

Making the Victim Whole: Torts Concepts

The typical objective of compensation in U.S. accident law is to make the victim whole.[5] The economic analog of this torts concept is to provide sufficient compensation to restore the individual's utility level. Although courts by their very nature provide *ex post* compensation after an injury has Occurred, it may also be the case that individual welfare is restored on an *ex ante* basis. Wage compensation for job risks promotes individuals' expected utility on an *ex ante* basis, whereas workers' compensation is a form of *ex post* compensation.

In the case of voluntary risks, the individual bearing the risk has received compensation on an *ex ante* basis sufficient to give him the same expected welfare level that he would have had if he had not borne the risk. In effect, his expected losses have been made whole *ex ante*. After the fact, depending on whether he is fortunate or not, he will be more than or less than whole.

For involuntary risks, it may Se desirable from an equity standpoint to restore the individual's utility level. In the absence of a mechanism for transferring money, one can in effect promote greater equity through in kind redistribution, in particular by regulating the risk to a lower level.

Consider two alternative situations in which it is desired to ensure an individual that he remain at his current expected utility level. In the first, one can transfer money and regulate the risk, and in the second no such cash transfers are feasible. The generator of the risk will be a Polluting firm, and for Simplicity there will be one party bearing the risk.

The firm produces its output q at a cost C(q) , where C, C′, and C" are all positive. Let v be the market price of the output. In addition, there is a unit cost G(z) to producing environmental quality z , where G′ and G″ are positive. One might view z as the level of safety so that $1-z$ is the risk. For concreteness, we will view z in probabilistic terms as the probability that an individual suffers an injury. The analysis can also be recast in terms of some non stochastic pollution level Or some other externality.

Individual consumption x will provide utility $U^1(x)$ if the individual is healthy and $u^2(x)$ if he is not, where $u^1(x) > u^2(x)$, $u^1 > u^1$) and both U^1_{xx} and $U^1_{xx} > 0$. There is a probability z that the healthy state 1 will prevail and a probability 1 z that the ill health state 2 will prevail. Let U^o be the baseline expected utility level in the absence of a risk, and for simplicity subsume the role of all income other than cash transfers as compensation for the risk into the functional forms of the utility functions. Then a given combination of risk Z and *ex ante* compensation y makes the individual whole if

$$(1) \qquad U^0 - zU^1(y) - (1-z)U^2(y) = 0 .$$

In the first situation let the government's role be to set up a mechanism to control the risk level and to establish the level of compensation.[6] The objective will be to establish a structure to maximize the polluting firm's profits subject to equation 1. In particular, it will

$$\max_{y,z} W = vq - C(q) - qG(z) - y + \lambda[xU^1(y) + (1-z)U^2(y) - U^o] ,$$

leading to the first order conditions

$$(2) \quad 0 = \frac{\partial W}{\partial y} = -1 + \lambda[zU_y^1 + (1-z)U_y^2] ,$$

and

$$(3) \quad 0 = \frac{\partial W}{\partial z} = -qG' + \lambda[U^1 - U^2] .$$

The analysis for compensation that makes the victim whole <u>ex post</u> rather than <u>ex ante</u> is quite similar, as the equity constraint is that

$$(1') \quad U^o = U^1(0) = U^2(y),$$

and the first order conditions are

$$(2') \quad 0 = \frac{\partial W}{\partial y} = -1 + \lambda[(1-z)U_y^2] ,$$

and

$$(3') \quad 0 = \frac{\partial W}{\partial z} = -qG' - \lambda U^2 .$$

If there is no potential for cash transfers, equation 1 (or 1') is omitted, and one must promote equity concerns through risk regulation alone. By sacrificing an additional policy instrument, this alternative will be less efficient. The shadow price λ on the equity constraint will be greater since the cost of promoting equality will have risen. In addition, a lower level of y will raise the level of z that is needed to restore the individual's utility. Indeed, if the individual was originally at a no risk position, the safety level z must equal 1 to restore his utility level.

In practice, the policy objective may be more limited to perhaps maximizing a weighted average of the firm's profits and individual utility- One can vary the value of λ to trace out the set of such Pareto optimal equity schemes.

In general, however, the elimination of compensation as a policy instrument, as occurs when we shift from voluntary to involuntary risks, has a twofold effect. First, it imposes an efficiency loss by taking us off the efficient frontier of equitable policies. Permitting the transfer of cash enables policies to be more efficient than if one must rely on risk regulation alone. Second, the increased reliance on risk regulation leads to higher values of safety z than would be optimal in the presence of compensation. Risk regulation will be used as a form of redistribution. As a result, there will be greater reliance on risk regulation as an equity promoting device for involuntary risks.

Making the Victim Whole: Efficient Insurance

When providing *ex post* compensation it will not typically be desirable to fully restore the victim's welfare. How large must a bequest to one's family be to give a victim of a fatal accident the same level of welfare as he would have had if he had lived? No amount is sufficient. Moreover , considerable compensation of this type will have only a small effect on the individual's welfare.

A more meaningful basis for conceptualizing appropriate compensation levels is to ascertain how much compensation the individual would choose to provide for himself if he were able to purchase insurance on an actuarially fair basis.[7] That compensation will not restore the pre-accident utility level but will instead equate the marginal utility of income in the two health states.

Providing such compensation by, for example, a workers' compensation system poses little problem in the case of involuntary risks since the actions of the individuals affected and the firm will be unaffected. For voluntary risks, government subsidized compensation will reduce the level of *ex ante* compensation that firms must provide individuals who bear the risk. Moreover, unless there is perfect merit rating, there will be a subsidy to risky activities, leading to increases in the supply of voluntary risks.

There is consequently a greater role that can be played by a penalty/compensation scheme for involuntary risks. Because the optimal penalty will generally be larger than the efficient level of *ex post* compensation, it will be desirable to separate the role of incentives and compensation by, for example, giving some of the penalty to the community at large that was exposed to the risk rather than giving all of the funds to the accident victims.

Making the Victim Whole: Lifetime Risk

Although the usual focus of equity concepts is on welfare or on income, here we will develop a different approach based on the level of the risk. In Particular, how can we approach the task of making the individual whole in risk terms?

One objective might be to restore the person to the no risk level after a risk exposure. This concept would often be unworkable since a person exposed to a positive risk cannot be exposed to some negative risk in the future except perhaps by reducing the risks of some other activities as well as that of the particular risk In question. More fundamentally, the appropriate policy objective is not zero risk but the efficient risk level. Here we will not delve into the issue of what particular policy objective should govern the determination of the optimal risk. It might be a benefit cost test, or an objective of eliminating all significant risks, or some other formulation. What we will take as given is that the optimal level of risk is not zero.

Consider the dose response relationships in Figure 1. Suppose the individual is exposed to an overly large risk AB over time, such as the cancer risks from radiation. At the time the risk is discovered to be too great, the risk is then regulated on an optimal basis thereafter, and the individual moves along BC . In contrast, if the risk had been at the optimal level since the beginning, the dose over time would have followed ADE .

In the case of immediate risks, such as most safety hazards, an individual who is not injured at point B will ‘have been exposed to greater previous safety risks, but his future risk of injury will be the same as if he moved along the optimal path DE . For longer term risks of a more cumulative nature, such as lead concentration in the blood or exposure to AIDS, optimal regulation along BC will lead to greater subsequent risk than if the individual could be moved to DE . The same amount of additional exposure will impose more additional risk when risks are cumulative.

In each instance, one could establish a rationale for reducing the risk sufficiently to move the individual from the dose path BC to DE . The rationale for moving along DE is to get to the optimal lifetime risk path from the path that involved an inordinately high risk exposure. Even though DC is optimal on a prospective basis, it exposes the individual to too great a lifetime risk, and in the case of cumulative hazards poses too great a future risk as well.

These concerns are likely to be more prominent for involuntary hazards for two reasons. First, in the case of voluntary risks individuals choose the dose so that it is not at all clear that society will wish to ensure them a low lifetime risk. Although ABC may not be optimal from a policy standpoint, if the risk was truly voluntary then the individual received compensation for the risk, any arguments about an unfair lifetime risk must be somewhat muted.

The second reason why involuntary risks are more likely to be affected is that risks that are involuntary often tend to be hidden hazards, such as dimly understood carcinogenic risks. The risks themselves may have been accepted voluntarily, but the level may have been uncertain. It is this hidden nature that contributes to the risk’s involuntary nature, since in the absence of good information individuals cannot engage in an informed market transaction to bear the risk. Risks that are hidden or are not well understood will tend to be discovered later, so that the AB segment will be longer. Consequently, for involuntary risks the gap between the current lifetime dose and the optimal lifetime dose will tend to be greater than for voluntary risks.

IV. Global Incentives and Regulatory Standards

In addition to considerations pertaining to benefit valuations and equity concerns, one must also take into account the structure of regulatory policies. In this section we will address the global incentives problems arising from the typical reliance on regulatory standards, and in Section 5 we will focus on inadequacies stemming from the character of the enforcement effort.

The Certainty Case

In the ideal situation one will obtain an efficient outcome if the firm must pay for society's valuation of the risk. For simplicity let s be the constant valuation per unit of risk imposed. In the case of an optimal risk penalty scheme, the firm will pay a fine s for each unit of risk so that the total penalty is the product of s , the level of output q , and the level of the risk 1Lz . For voluntary risks, this penalty will, in effect, be assessed by the market as s will be the price of the risk that is, for example, reflected in the wage premiums workers receive for hazardous jobs.

In such a situation the firm's profits are given by

$$(4) \quad \max_{q,z} \pi_a = vq - C(q) - qG(z) - sq(1-z) ,$$

where the firm picks q and z to satisfy

$$(5) \quad C' = v - G(z) - s(1-z) ,$$

and

$$(6) \quad G' = s .$$

It is straightforward to replicate the risk levels achievable under penalties with a suitably designed standards program. One simply equates the marginal benefits of risk reduction, s , with the marginal costs, G′ . Optimal standards of this type produce the outcome under equation 6 through direct control of the risk rather than through a change in the incentives facing the firm. Below we will denote this optimal risk level by Z^*_a .

Even though the risk level achieved is the same as under a penalty system, the global incentives are quite different. In particular, with standards the firm's profits are given by

$$(7) \quad \max_q \pi_b = vq - C(q) - qG(z^*_a) ,$$

where the firm picks q so that

$$(8) \quad C' = v - G(z^*_a) .$$

Two observations are noteworthy. First, the firm reaps more profits under the standards system given by equation 7 than it does under the penalty scheme reflected in equation 4. A firm facing regulatory standards receives all of the pollution up to the level of the standard free. As a consequence, even though standards can be used to set efficient local incentives for controlling risk, the global incentives will not be efficient. There will be too much entry into risky industries with standards that equate the marginal benefits and marginal costs of risk reduction.

A second difficulty is that the level of output selected under standards will be too great. The optimal output with penalties satisfies equation 5, which includes a term s(l z) that captures the effect of additional output on the level of the risk. There is no such term in equation 8, leading to a higher level of output.

If the regulatory mode must be standards, then following the conventional guidelines of equating marginal benefits and marginal costs will lead to too much entry and too high a level of output in the risky industry. Setting the standard at a more stringent level than the usual norm for efficiency will alleviate each of these problems. Higher levels of safety z will raise the costs $G(z)$, reducing the incentives to enter the industry. Similarly, this higher value will reduce the output level selected.

No standard can replicate the outcome of an optimal penalty scheme with respect to the global incentives for entry, the incentives for picking the output level, and the level of the risk. By overregulating the risk we sacrifice some efficiency with regard to the final objective in an effort to foster the first two objectives. These concerns have been generally ignored in discussions of regulatory standards.

The Uncertainty Case

Because standards set at the “marginal benefits equal marginal costs" norm do not force firms to take into account the full economic effects of their actions, there will also be a tendency for firms to incur too many risks under a conventional standard setting scheme. Consider a two period model, where we will index each of the earlier variables by the appropriate time period subscript i, where $i=1, 2$. Let β be the discount factor, or the inverse of one plus the interest rate. Suppose that in the initial period the firm’s technology is believed to be risk free so that there are no risk control costs, but there is a chance p that the firm’s emissions will be found to pose a nonzero risk z_2 per unit output that the firm can reduce through appropriate expenditures.

Under a penalty system, the firm's profits over the two periods are given by

$$\max_{q_1} \pi_a = [vq_1 - C(q_1)][1 + \beta(1-p)] + \beta p \max_{q_2, z_2} [vq_2 - C(q_2) - q_2 G(z_2) - sq_2(1-z_2)] \; .$$

Under a standards system that establishes the same optimal level of z_2^* that would be selected with penalties, the firm's problem is to

$$(9) \quad \max_{q_1} \pi_b = [vq_1 - C(q_1)][1 + \beta(1-p)] + \beta p \max_{q_2} [vq_2 - C(q_2) - q_2 G(z_2^*)] \; .$$

For much the same reason as above for the certainty case, π_b will exceed π_a . When a firm is comparing technologies with different risks of being discovered as being harmful, it will have too great an incentive to pursue a technology that poses a risk of being found to be harmful in the future. The firm is being given a future subsidy of $sq_2^*(1-z_2^*)$ with respect to such technologies, where q_2^* is the optimal value of q_2 in equation 9.

There is a discounted expected subsidy in the initial period of $\beta psq_2^*(1-z_2^*)$. The bias toward the choice of risky technologies increases with the level of the discount factor, the probability that the firm's actions will be found to be harmful, the severity s of the harm, the scale of the output, and the unit probability tht the harm occurs (e.g., the worker injury rate per unit output).

In terms of the financial incentives created, standards that are set based on the usual norms will lead to a series of biases that all promote the selection of technologies that pose potentially major risks. These biases can be diminished by setting the standard at an excessively stringent level. Such compromises are not required for penalties policies or for fully voluntary risks because the global incentives are not distorted in those cases.

V. Enforcement Related Concerns

As was noted in Section 3, risks that are difficult to monitor tend to fall into the involuntary class because in the absence of knowledge of the risk one cannot undertake the market transactions that convert risks into being voluntary in nature. These hidden risks impede not only market processes but enforcement of government regulations as well. In the case of job risks, safety hazards can be identified fairly readily by an OSHA imspector, whereas monitoring of health risks is a very time consuming process. Typical safety inspections of a diverse set of potential safety risks average 10 hours per inspection, whereas the much more narrowly focused health inspections average 33 hours.[8] Similarly, in the environmental area, monitoring of toxic wastes is very spotty, in large part because the pollution is much more clandestine than, for example, emissions from a smokestack. Compliance rates for hazardous and toxic substance regulations are only 20 percent, largely because of these monitoring problems.[9]

The low probability of detection of involuntary risks could be offset by making the penalty for violations sufficiently large. Doing so in situations where the probability of discovery is low would lead to very substantial penalties that might appear unfair and would not be optimal if firms were risk average. At present, penalty schemes are not boosted in situations where the detection probability is low. In the absence of such variation, one can augment the incentives for controlling risks by making the standard more stringent. Doing so will increase the probability of detecting a violation and raise the extent of the violations, thus boosting the overall penalty.

A final consideration is that such overregulation of involuntary risks may be warranted to the extent that society wishes to punish wrongdoers. Ideally, we would like to encourage a system of involuntary restraint. Indeed, much of our social system, ranging from the enforcement of income tax requirements to automobile driving practices, is based on the assumption that individuals will exercise restraint voluntarily. Increases in the stringency of the regulation consequently may be warranted to the extent that the failure to control the risk voluntarily poses a loss on society above and beyond the expected health effects.

VI. Conclusion

The conclusion that involuntary risks should be regulated more stringently is not surprising. Such notions have often been suggested, but never fully articulated.

What is striking is the diversity of concerns that support this distinction. Involuntary risks are likely to be borne by a different mix of individuals with higher valuations of health and less well equipped to handle the risks, and these risks are more likely to be valued highly in their own right. Equity concerns, particularly our concept of a fair lifetime risk, are also more likely to come to bear. Because government regulations rely on standards rather than penalties and because these standards are not fully enforced, there are additional reasons to tighten regulations for involuntary risks.

In terms of policy practice, such distinctions have only been recognized informally, at best. There is often greater public pressure for regulation of involuntary risks. These efforts primarily influence the regulatory targets. The regulatory agencies' standard setting process has not, however, incorporated recognition of the degree of volition regarding the bearing of the risk. Rather than being a second order concern, this distinction may be most fundamental in determining whether a risk should be regulated and to what extent.

Footnotes

[1] Most discussions of the properties of transactions for risks or other externalities make at least implicit use of the Coase Theorem. See Coase (1960).

[2] A similar distinction was employed originally by Mishan (1971) who does not delve into the class of issues considered in this paper.

[3] This discussion and the empirical evidence contained therein is based on Chapter 6 of Viscusi (1983). See Fischhoff et al. for a review of the underlying evidence on risk perception.

[4] The estimates of the heterogeneity of the value of life presented in Viscusi (1983) are also consistent with evidence for samples concentrating on particular risk ranges, such as Thaler and Rosen’s (1976) results for workers in high risk jobs and Viscusi's earlier analysis of workers in more moderate risk jobs.

[5] Calabresi (1970) provides a full discussion of these concepts.

[6] Although the firm’s choice of the output level will be treated as given, alternatively one can view it as being determined by a reaction function whereby the firm picks q to maximize profits given the level of y and z. By the envelope theorem, we can ignore all derivatives of q with respect to the policy choice variables in the analysis below.

[7] This approach and the general result that one should equate the marginal utility across health states, is incorporated in the models of Cook and Graham (1977), Spence (1977), and Viscusi (1979, 1980).

[8] These figures are for FY 1983 and are based on unpublished OSHA printouts.

[9]See W. Kip Viscusi, Richard Zeckhauser, and Peter Schuck, “Procedures for Setting Optimal Penalties for Environmental Regulation,” Report to the U.S. Environmental Agency (1986) Chapter 3.

References

Calabresi, Guido, The Costs of Accidents: A Legal and Economic Analysis (New Haven: Yale University Press, 1970).

Coase, Ronald, “The Problem of Social Cost,” Journal of Law and Economics, Vol. 3 (1960).

Cook, Philip, and Daniel Graham, “The Demand for Insurance and Protection: The Case of Irreplaceable Commodities ," Quarterly Journal of Economics, Vol. 91 (Feb. 1977), pp. 143 154.

Fischhoff, Baruch, Sarah Lichtenstein, Paul Slovic, Stephen Derby, and Ralph Keeney, Acceptable Risk (Cambridge, U.K. : Cambridge University Press, 1981).

Mishan, Ezra J., “Evaluation of Life and Limb: A Theoretical Approach,” Journal of Political Economy, Vol. 79 (1971), pp. 706 738.

Spence, Michael, "Consumer Misperceptions, Product Failure and Producer Liability,” Review of Economic Studies, Vol. 44 (1977), pp. 561 572.

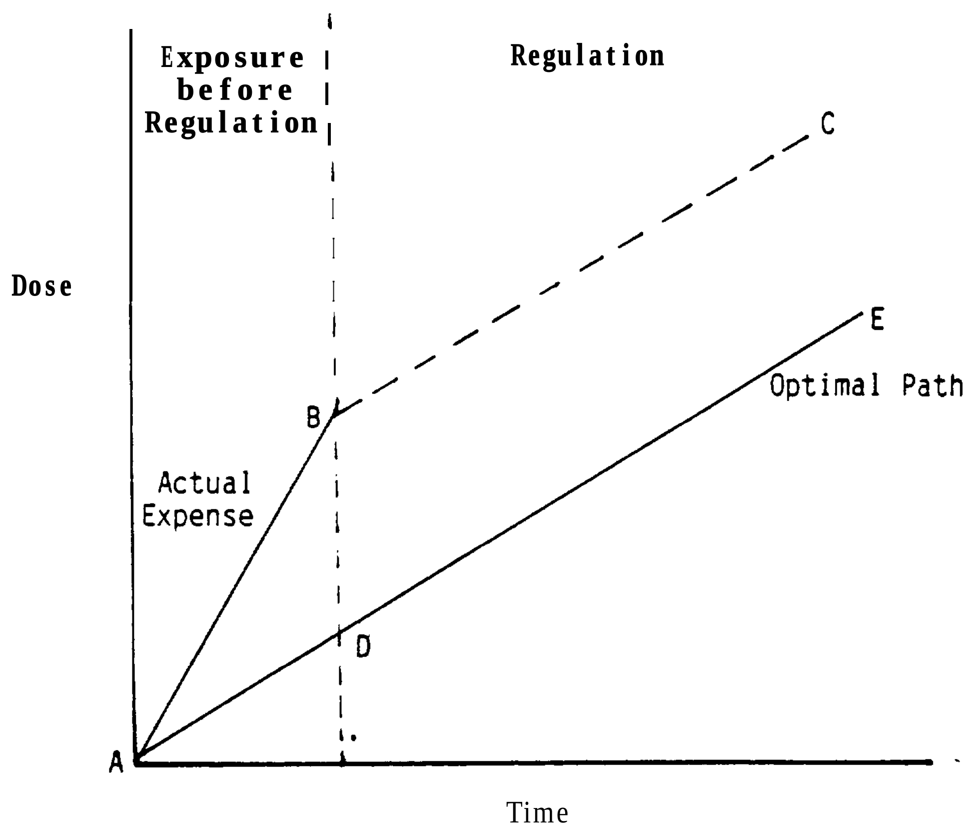
Thaler, Richard, and Sherwin Rosen, “The Value of Saving a Life: Evidence from the Labor Market.” in N. Terleckyj, ed., Household Production and Consumption (New York: Columbia University Press, 1976).

Viscusi, W. Kip, Employment Hazards: An Investigation of Market Performance (Cambridge: Harvard University Press, 1979) .

“Imperfect Job Risk Information and Optimal Workmen’s Compensation Benefits,” Journal of Public Economics, Vol. 14, No. 3 (l980), pp. 319 337.

Risk by Choice: Regulating Health and Safety in the Workplace (Cambridge: Harvard University Press, 1983).

Figure 1


Exposure before Regulation
Regulation
Dose
C
E
Optimal Path
B
Actual Expense
D
A
Time